TOWARDS AN AUTONOMOUS COMPUTATIONALLY INTELLIGENT COGNITIVE SYSTEM

John Taylor (KCL)
Department of Mathematics
King's College London, UK
john.g.taylor@kcl.ac.uk
http://www.kcl.ac.uk/research/cns/cns.html

KING'S College LONDON

University of London

WHAT?

- Autonomy = ability to move validly under own steam (possessed by many animals)
- autonomy as evolutionary tree
- Possession of autonomy ≠ intelligence
- eg persistent criminals/ schizophrenics/….
- Add intelligence (computationally created)
- But dangerous => wars/mayhem (War of Worlds), suicide bombers, etc, etc

HOW?

- Lessons from existing solution – humans
- Brain arguably subtlest ‘machine’ in universe (is conscious, not possessed by universe – insanely produces black holes)
- Need emotional value/empathy (‘love’?)
- => Need careful guidance of development of ‘conscious’ machine with ‘guided creative fun’ (learns value of NO)
- As should be in human upbringing

=> BRAIN GUIDANCE

- Look for brain principles (intelligence & emotion)
- Propose highest level control: attention
- Biased by emotion valuations of world
- Internal rules developed with biasing valuations
- Needs to be highly adaptive (STDP, reinforcement & error learning schemes)
- Use to build brain-based general cognitive computationally intelligent architecture

CONTENTS

1. NATURE OF COGNITION

- Cognition defined as:
- “That operation of the mind by which one becomes aware of objects of thought or perception; it includes all aspects of perceiving, thinking, and remembering”
- “Mental functions such as the ability to think, reason, and remember”
- “High level functions carried out by the human brain, including comprehension and use of speech, visual perception and construction, calculation ability, attention (information processing), memory, and executive functions such as planning, problem-solving, and self-monitoring”
- => Cognition complex!

How to look at cognition

- Reasoning, planning and self-monitoring: crucial components of cognition
- Leave out speech: animals can reason => look at non-linguistic cognition
- Approaches for reasoning, planning and self-monitoring:
- a) Symbolic: logical inference on language structures
- b) Probabilistic: cognition = probabilistic inference
- c) Connectionist: how can inference be obtained from neural network structures at sub-symbolic level?
- Look at neural structures as most relevant here (can best relate to brain processing methods)

Problems for Autonomous Machines

- **Problems of machines presently:**

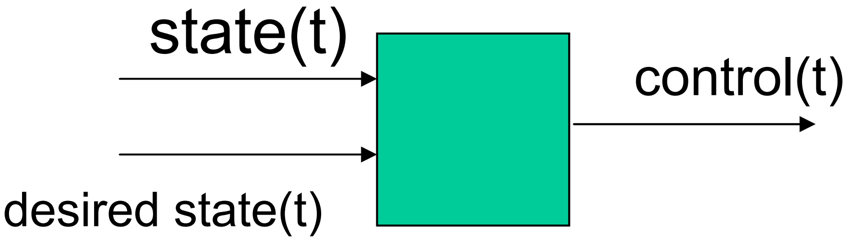
1) **Scalability** (many sensors, real time)
2) **Context awareness** (peripheral sensitivity)
3) **Robustness** (against damage/loss)
4) **Autonomy & self-management** (stay alive)
5) **User adaptability** (varying profiles)
6) **Fast computation** (rapid guidance)

Essential structures

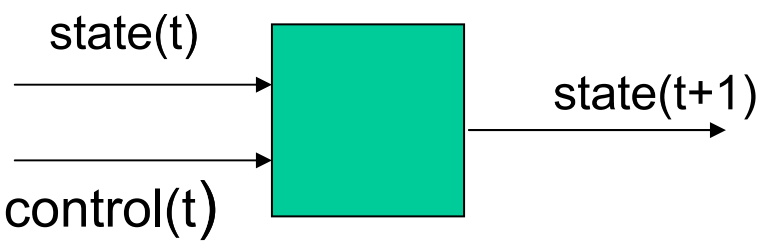
- Crucial components = internal models of world
- Forward models (encode causality of the world)
- Working memory modules (imagine the world)
- Inverse model controller IMC (generates desired control actions)
- Filter control system (filters world into ‘to be attended to’ + ‘rest = distracters’)
- FM/IMC by control theory (no WM or attention)

Control Model Components I

Inverse model controller (IMC)

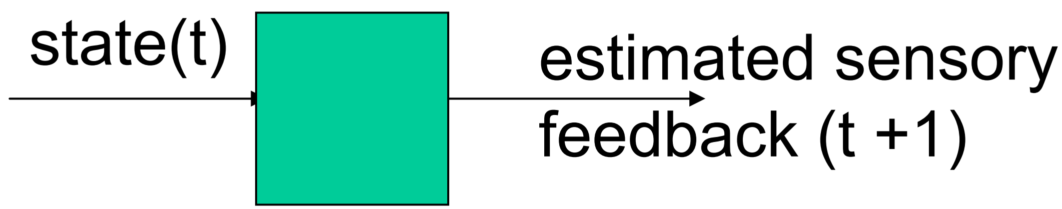


Forward model/ observer

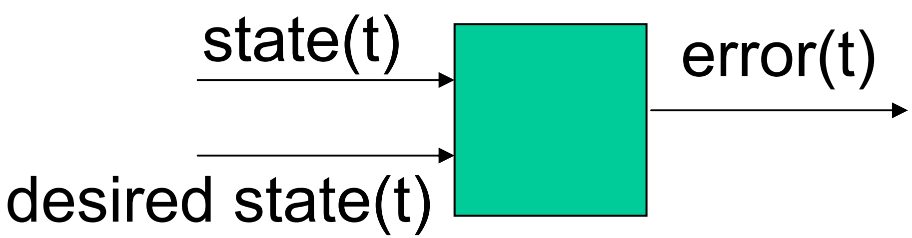


Control Model Components II

Forward Output Model



Error Monitor Module



To be used in error-based learning (as in Cb)

Solution to Autonomous Machine Problems: Learn to Pay Attention!

- **Attention solves:**

1) Scalability: filter out distracters

2) Context awareness: attend to important

3) Robustness: use neural multiplexing/pop

4) Autonomy: determine by own goal structure

5) User adaptability: by training to a user

6) Speed: by hardware nanotechnology

Extra Components for Cognitive Powers

- Show later how can model reasoning powers of animals by dedicated coupled FM/IMC/WM systems
- Used to generate & transform ‘world in head’ to achieve executive functions
- Attention crucial in complex world
- Need good model of attention/ learning/ WM/ executive powers/ STM/ LTM: all powers possessed by humans

2. COGNITIVE BRAIN-BASED MACHINE PROJECTS

- GNOSY ‘Cognitive Robot’ (EC IST Project Oct 2004 – ‘07)
- MATHESIS ‘Learning Others’ Actions’ (EC IST Project Feb 2006-’09)
- ‘Attending to the World’ (EPSRC Project March 2004 – 2007)
- ‘Analysing Attention<->Emotion’ (BBSRC Project 2005-2008)
- ‘Modelling Emotion’ (EC HUMAINE NoE, 2005-2008)

Partners/People Involved

- KCL: JG Taylor, N Taylor, M Hartley, C Panchev, N Fragopanagos, N Korsten (CNS)
- FORTH: P Trahanias/ S Kasderidis (CS)
- UTUB: H-P Mallot/ W Huebner (Vision)
- UGDIST: P Morasso, V Moran (Motor Actions)
- ZENON: C Emmanouilidis/VSpais (Robotics)
- Oxford (A Nobre), Bangor (K Shapiro/J Raymond), Birkbeck (M Eimer)
- FORTH (H Savaki), CNRS Paris (J Nadel, J Fagard), U of Bologna (C Galletti, P Fattori),
- HUMAINE (NoE, R Cowie + 150 more)

Focus of GNOSYS

1) Develop percepts/ concepts/ rewarded-goals/ reasoning/ abstraction
2) Learn to perform goal-directed tasks
3) Learn in novel environments
4) Reasoning by forward model
5) Globally integrated system
6) Employ various memory types (STM/ LTM/ iconic/ associative/motor)
7) Interdisciplinary: Comp vision/ Cog Sci/ Cog NSci/ Robotics/ Control/ AI/ Maths

How GNOSYS Works

V V

ANN Adaptive Streams (Concepts/ Goals/ Attention/ Rewards/ Values/ Forward Models learnt as NN predictors)

| ↔ |

| ↔ |

| ↔ |

| ↔ |

| ↔ |

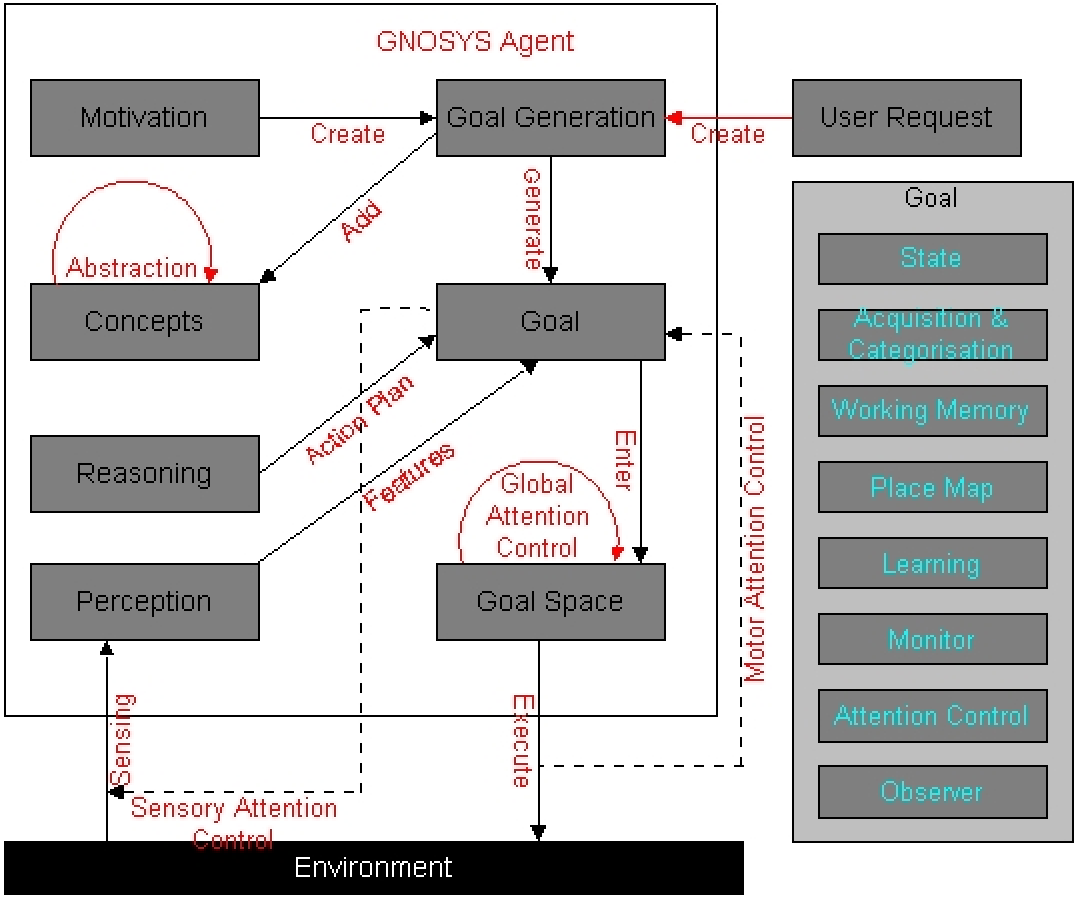
Symbolic Control Threads (several components)

Linguistic Connections (Words/Fuzzy rules/ Symbolisation)

GNOSYS Reasoning Domains/Environments

- **Three levels of environment**
- **Level 1: Learn shapes/colours; move & touch: [2]-D objects**
- **Powers: Concept/Attn/ Goals/Actions on objects/Salience of objects in environment**
- **Level 2: [3]-D objects & actions: pick-up, stack, learn new objects**
- **Powers: ibid/ manipulate to achieve goals**
- **Level 3: Hierarchy of objects; run virtual object/ action sequences to achieve goals**
- **Powers: Reasoning/ novel objects/ actions**

The Hybrid GNOSYS Brain (S Kasderidis, FORTH)



EC Cognitive Systems (Unit 5)

- ‘Focus is on research into ways of endowing artificial systems with high-level cognitive capabilities, typically perception, understanding, learning, knowledge representation and deliberation, thus advancing enabling technologies for scene interpretation, natural language understanding, automated reasoning and problem-solving, robotics and automation, that are relevant for dealing with complex real-world systems. It aims at systems that develop their reasoning, planning and communication faculties through grounding in interactive and collaborative environments, which are part of, or connected to the real world.
- These systems are expected to exhibit appropriate degrees of autonomy and also to learn through "social" interaction among themselves and/or through human-agent cooperation; in a longer term perspective, research will explore models for cognitive traits such as affect, consciousness or theory of mind.’

How To Assess?

- Ambitious, even mentions consciousness and theory of mind.
- Similar ambition in other new adventures in cognitive research: Brain Sciences Institute (BSI) in Tokyo, BICA (USA) – emphasise brain basis/use of brain guidance
- BSI made good progress towards its aims
- High ambition not => negative reaction
- But need careful assessment of projects & results
- Need to realise some goals harder than initially thought
- Eg consciousness!

The EC Cognitive Systems Projects

- The 23 ACS projects listed below: (in alphabetical order):
- 1) BACS: Bayesian Approach to Cognitive Systems
- 2) CASBLIP: Cognitive Aid System for Blind People
- 3) CLASS: Cognitive-Level Annotation using Latent Statistical Structure
- 4) COSPAL: Cognitive Systems using Perception-Action Learning
- 5) COSY: Cognitive Systems for Cognitive Assistants
- 6) DECISIONS-IN-MOTION: Neural Decision-Making in Motion
- 7) DIRAC: Detection and Identification of Rare Audio-visual Cues
- 8) eTRIMS: eTraining for Interpreting Images of Man Made Scenes
- 9) euCOGNITION: European Network for the Advancement of Artificial Cognitive Systems
- 10) GNOSYS: An Abstraction Architecture for Cognitive Agents
- 11) HERMES: Human-Expressive Representations of Motion and their Evaluation in Sequence

Cognitive Systems II

- **12) ICEA: Integrating Cognition, Emotion and Autonomy**
- **13) JAST: Joint-Action Science and Technology**
- **14) MACS: Multisensory Autonomous Cognitive Systems**
- **15) MATHESIS: Observational Learning in Cognitive Agents**
- **16) Mind RACES: from Reactive to Anticipatory Cognitive Embodied Systems**
- **17) PACO-PLUS: Perception, Action and Cognition through Learning of Object-Action Complexes**
- **18) PASCAL: Pattern Analysis, Statistical Modelling and Computational Learning**
- **19) POP: Perception On Purpose**
- **20) RASCALLI: Responsive Artificial Situated Cognitive Agents Living and Learning on the Internet**
- **21) ROBOT-CUB: Robotic Open-architecture Technology for Cognition, Understanding and Behaviours**
- **22) SENSOPAC: SENSOrimotor structuring of Perception and Action for emerging Cognition**
- **23) SPARK:Spatial-temporal patterns for action-oriented perception in roving robots**

Assessment approach

- Gather the projects together under several headings, emphasisng various approaches to cognition::
  - Embodiment driven (# 13, 21)
  - Applications-driven (# 2, 5, 7, 8, 20)
  - Machine-intelligence driven (# 1, 3, 18)
  - Neural-based (# 15)
  - Cognitive science based (symbolic)
  - Hybrid (# 10)
  - Dynamic systems (# 23)
- Various approaches need not be most effective to achieve breakthrough in creation of autonomous cognitive machine.
- Need general model of human cognition to properly assess projects viability

Other Groups

- Numerous other groups: AIBO (Sony), AZIMO (Honda), Darwin 1–N (G Edelman, La Jolla), K Kawamura (Developmental Robotics), ATR Lab, Kyoto (M Kawato), COG Lab (R Brooks, MIT), 'Conscious Robot' (O Holland, Essex) + many more
- Use of HDP (BEP-based) very effective in creating motor control of walking robots (Kawato/Doya)
- 'Imitating robots' important area (A Billard)
- Progress across many fronts

3. BRAIN-BASED PRINCIPLES FOR COGNITION

- Look for highest-order control system in the brain
- Arguably=attention (sensory & motor)
- But even higher = consciousness
- Need to go through attention to reach neural functionality for consciousness
- => attend most carefully to attention!

Attention & Goals

- Attention both local & global
- Local: Controversy over synaptic effects (eg Grossberg/Deco & Rolls/Itti & Koch)
- Global (Ling & Carrasco, 2005):
  Contrast gain (endogenous)
  Output gain (exogenous)
- No sign of additive feedback?
- Detailed simulation (NT/JGT) => σ-π

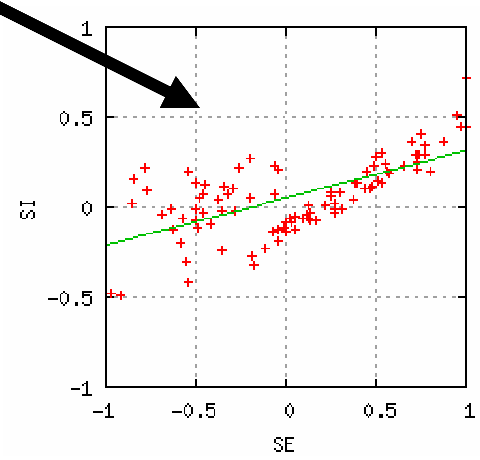
Local Mechanism of Attention:

- **3 mechanisms: output gain/added feedback/contrast gain**
- **Contrast gain: ↑weight of attended input**
- **Simulations (MR/NF/NT/JGT; Reynolds; use contrast gain):**

  Good fits to V2/V4 single cell monkey data => only contrast gain

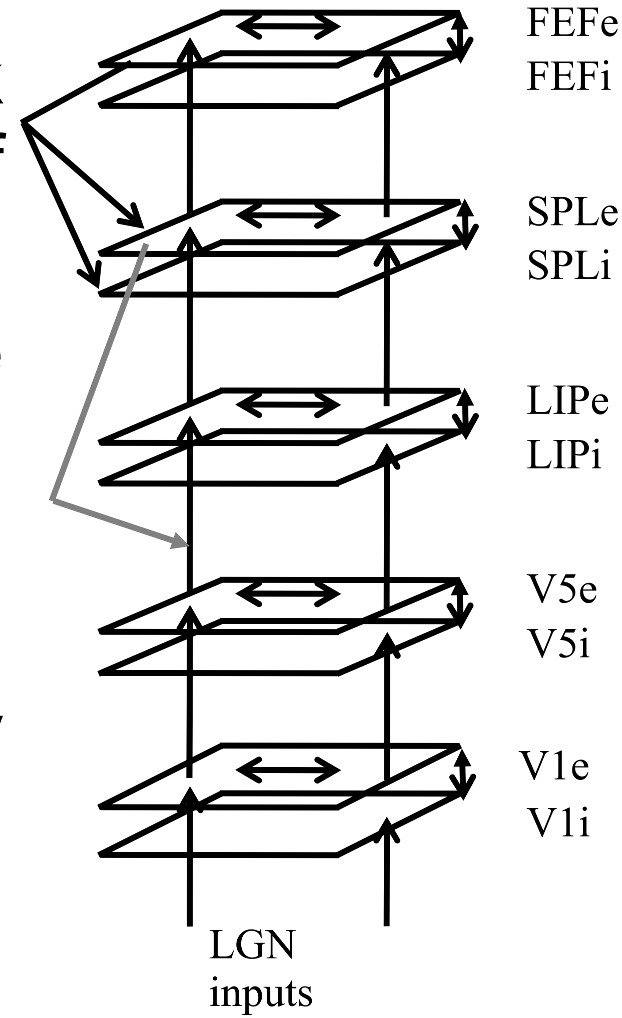
Simulation Results (NT/JGT/MH: IJCNN05, BICS06)

- Only contrast gain => matching graphs of experiments of Reynolds et al 1999
- Additive => 2 groups of neurons (attend probe/attend reference
- Not same regression lines as for original line
- => only contrast gain



Global Sigma-Pi Simulation

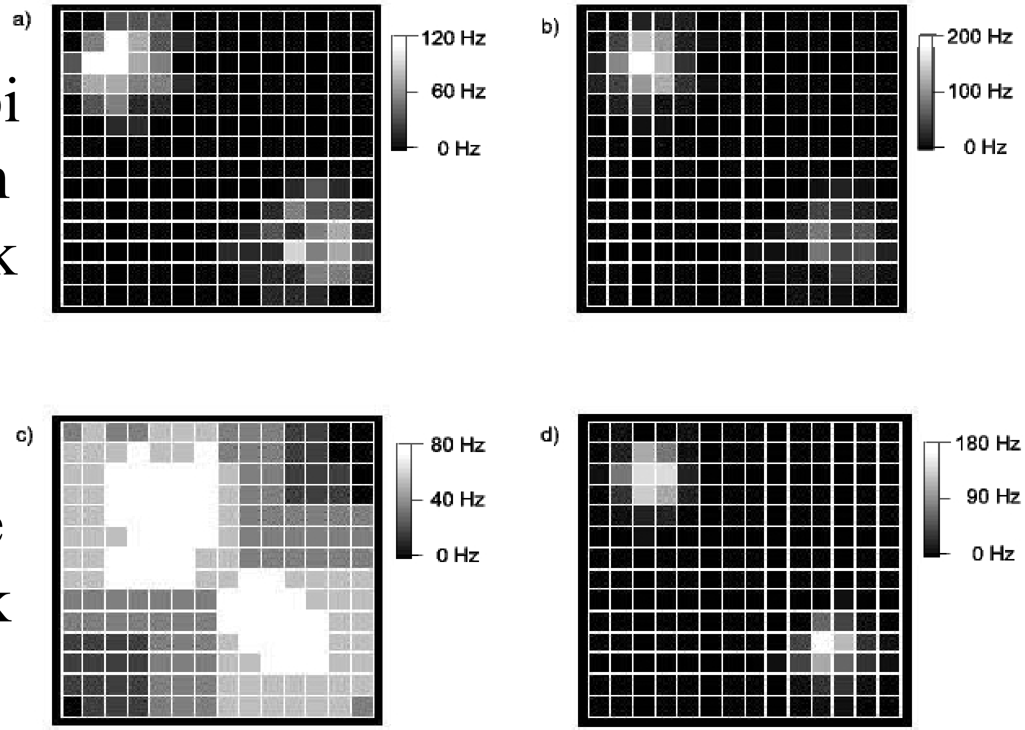
- Sigma-pi attention feedback investigated simple model of spatial attention
- Model based on dorsal route of visual cortex, pareital areas & frontal eye field(FEF)
- Upward flow refines spatial representation via inhibitory layers



Sigma-Pi Simulation (Results)

Sigma-pi attention feedback

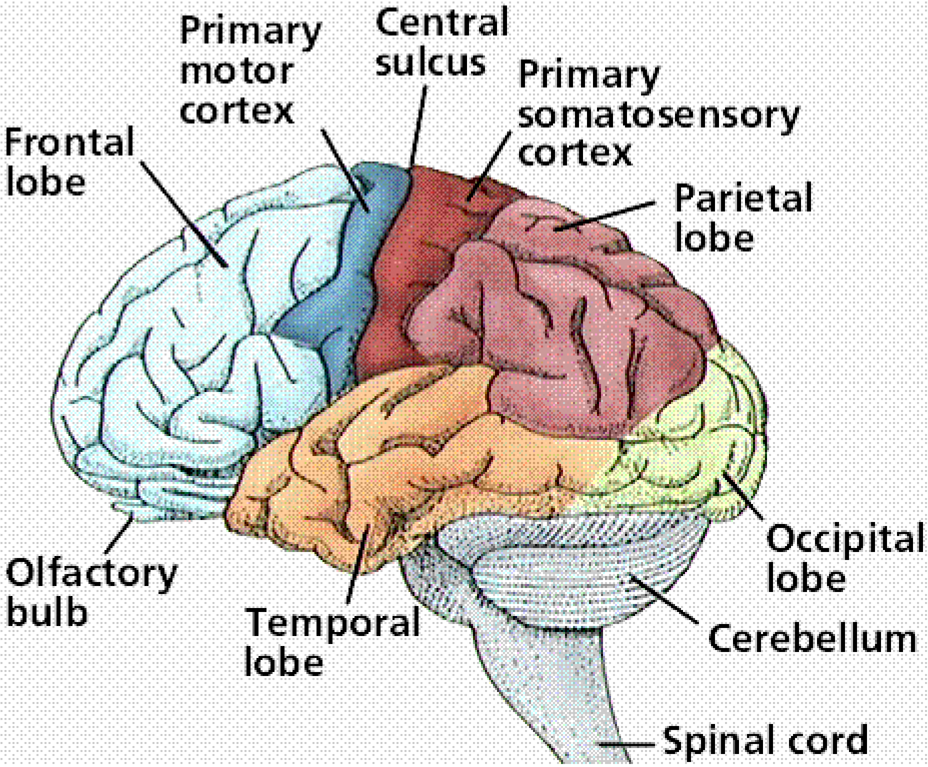
Additive feedback



Activity in LIPe neurons all figures show averaged firing rates (50ms). Figures a) & b) have sigma-pi weights, goal loc. is top left. Figures c) & d) indicate LIPe rates without sigma-pi weights. In both case the goal loc. is held in FEF

Global Level: The Human Brain

A very complex system:
10^{11} neurons, 10^{14} connections
Many hypercolumns & modules
Subtle functionality
Subtle pattern emergence
->High level functionality

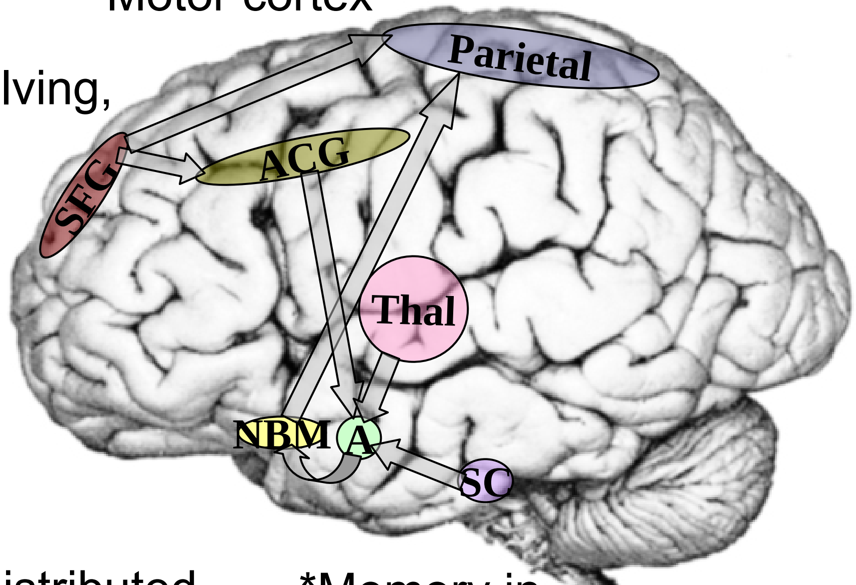


Overview of Brain Parcellation of Function

*Action in
Motor cortex

*Attention (controlled by parietal)

*Problem-solving,
Creativity:
Executive fn



*Perception
(in Visual, Somatic
& Auditory cortices

*Language is distributed
In the brain
*Consciousness siting is
unknown – also distributed?
Most crucial: parietal lobes

*Memory in
Hippocampus

Detailed & Separate Brain Functions

- Brain imaging of components of cognition show separate brain networks
- Objects: face in fusiform area (temporal lobe); others in nearby TL regions
- Attention: controlled by parietal lobes
- Response: in motor areas
- Response guidance (rewards): OBFC/Amygdala/Nucleus Accumbens

Attention as Global Control

- **‘ATTENTION = SELECTION OF PART OF SCENE FOR ANALYSIS’**
  (filter on input: covert - beyond James)

- **AMPLIFICATION OF ATTENDED + INHIBITION OF DISTRACTORS**
  (in sensory & motor cortices, & higher sites)

- **TWO PARTS:**
  1) ATTENDED COMPONENT
  2) ATTN CONTROL SIGNAL GENERATION

- **INCREASED ACTIVITY LEVEL WHEN ATTENTION DIRECTED TO SENSORY INPUT**

  (from EEG/PET/ FMRI/ MEG and single cell results)

- **MODULATION OF V4 CELL RESPONSE**
  (Maunsell et al, J NSci 19:431, 1999)

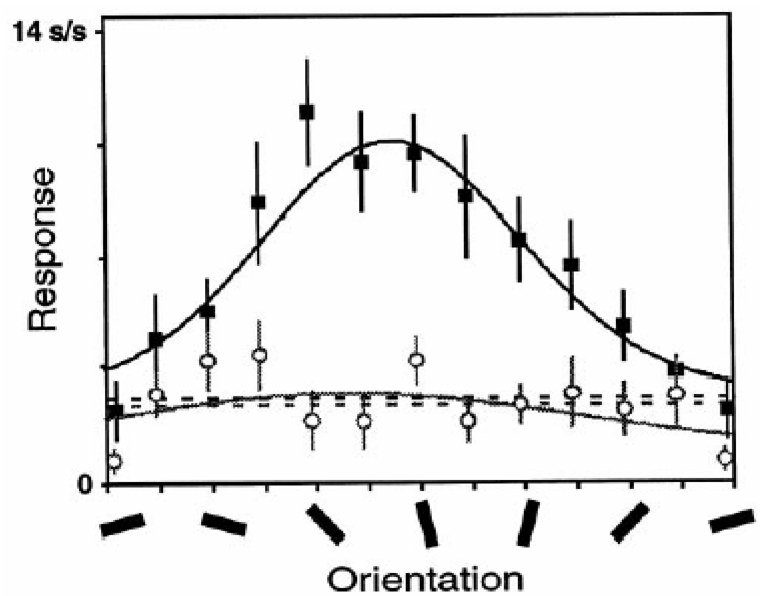


FIG. 2. Data from one V4 cell showing enhanced responses in the attended mode (*black*) relative to the unattended mode (*gray*)

- **ATTENTION MOVEMENT BY NETWORK OF BRAIN SITES:**
  - ∗ **PARIETAL (control)**
  - ∗ **FRONTAL (control)**

- **Shifting Attention Network (Corbetta, PNAS 95:831, 1998)**

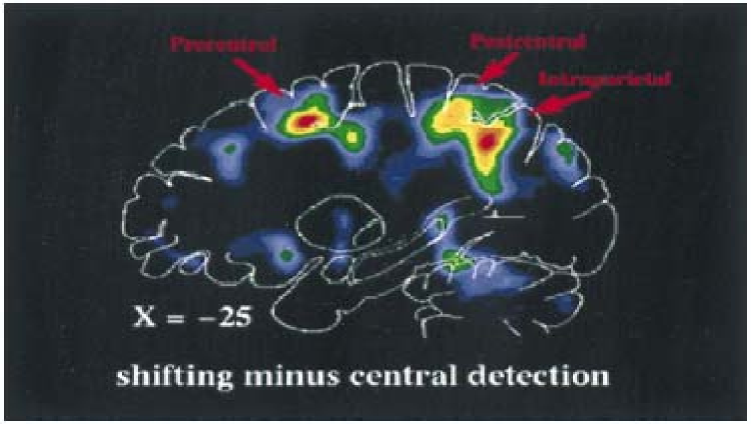
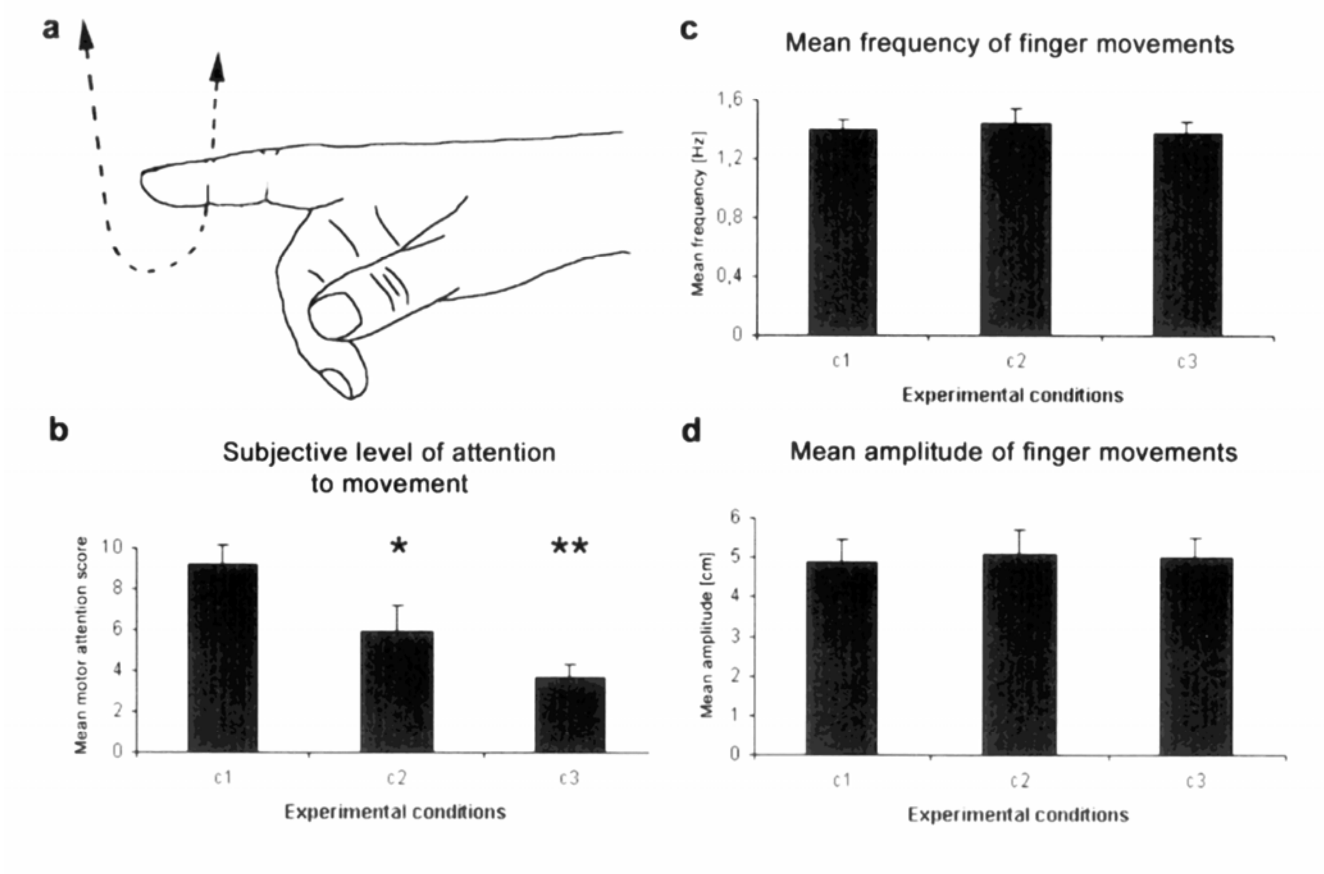


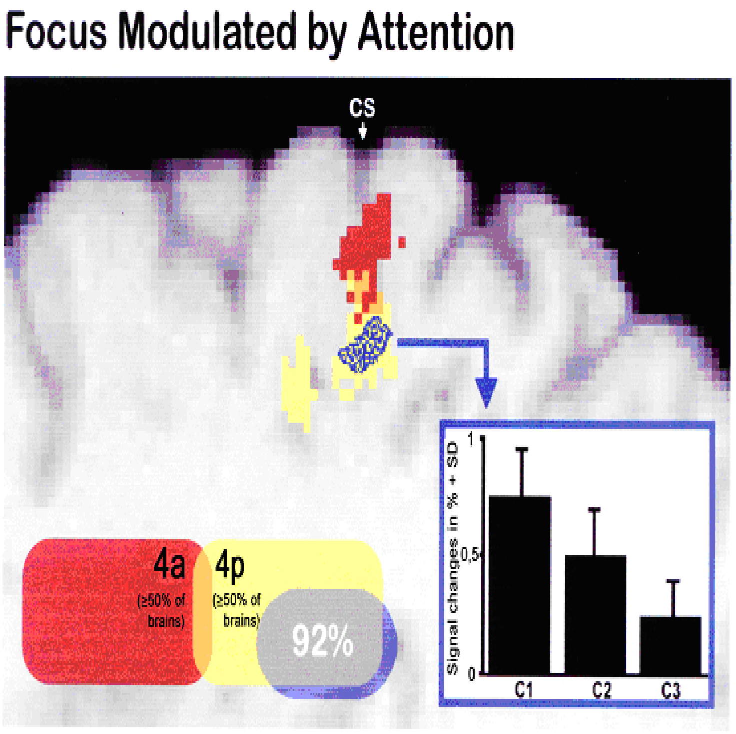
FIG. 1. Sagittal PET section, 25 mm left of midline, of group-averaged subtraction image between shifting-attention and central-detection tasks.

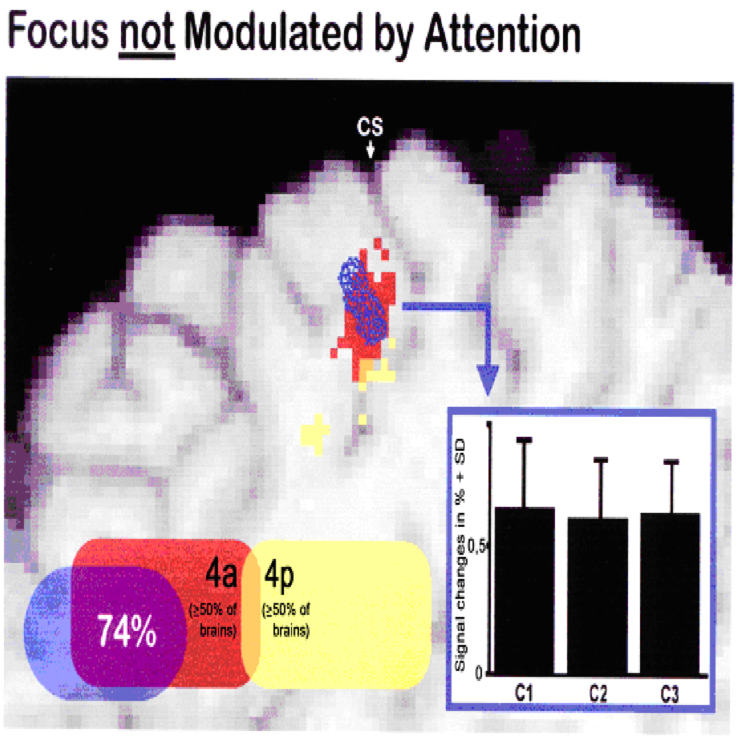
Paradigm for Motor Attention Study (Binkofski et al, J Neurophysiol 2002)



- **INCREASED ACTIVITY IN MOTOR CORTICAL SITES FOR ATTENTION TO RESPONSE**

  (Binkofski et al, J Neurophysiol 2002)





Global Attention Control:

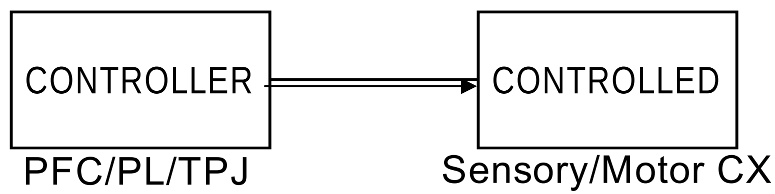
- ***OVERALL:*** **ATTENTION MOVEMENT INVOLVES BRAIN SITES WITH 2 DIFFERENT FUNCTIONS:**

  ∗ AMPLIFICATION/DECREASE OF SENSORY INPUT

  (in sensory & motor cortices)

  ∗ CREATION OF CONTROL SIGNALS TO DO THIS

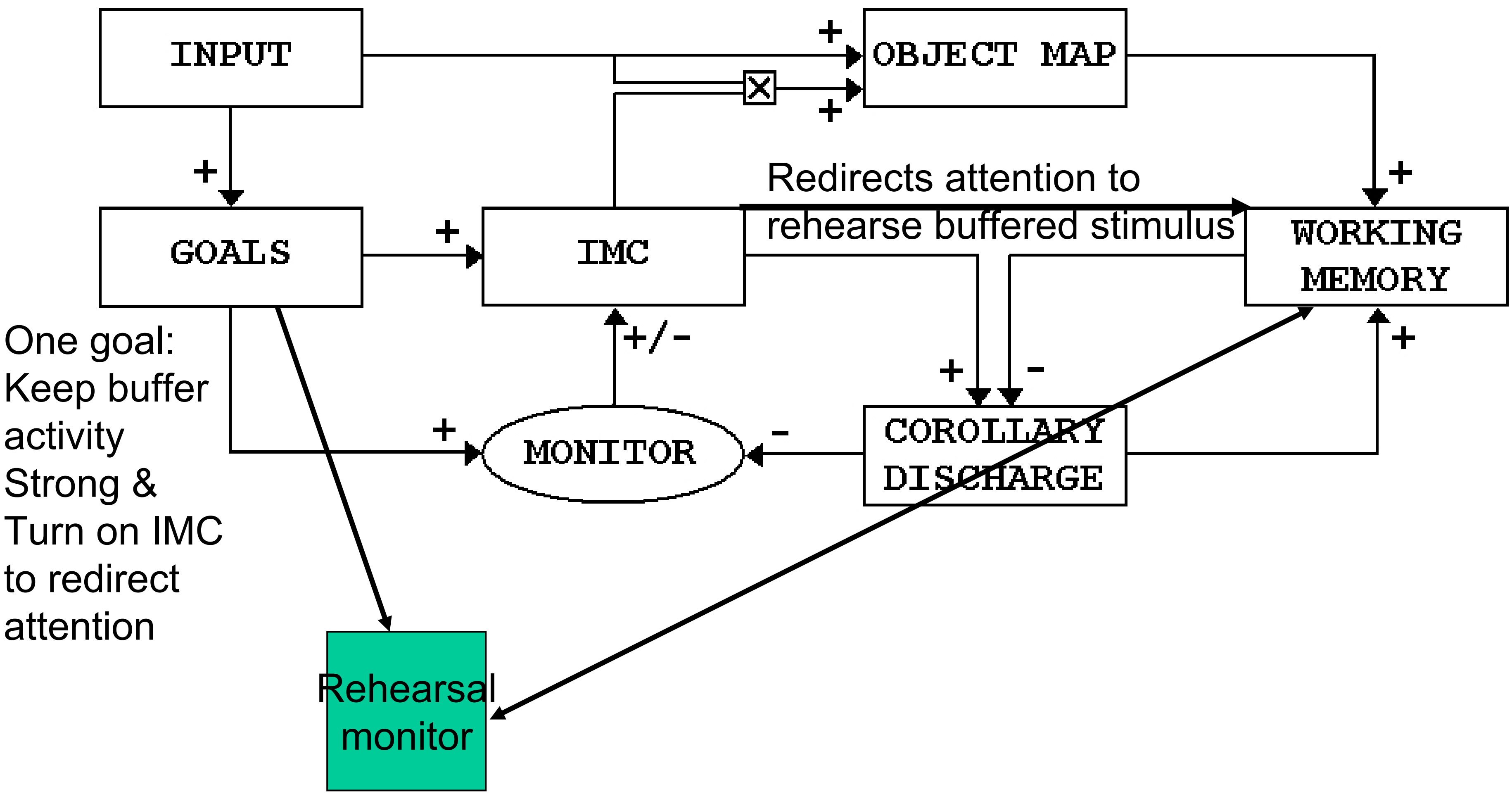
  (in parietal & frontal cortices):



- **EXPECT SITES IN BRAIN WITH SPECIFIC FUNCTIONS TO ACHIEVE THIS CONTROL**

  (goals, monitors/errors, feedback signal control generators)

Extended CODAM model



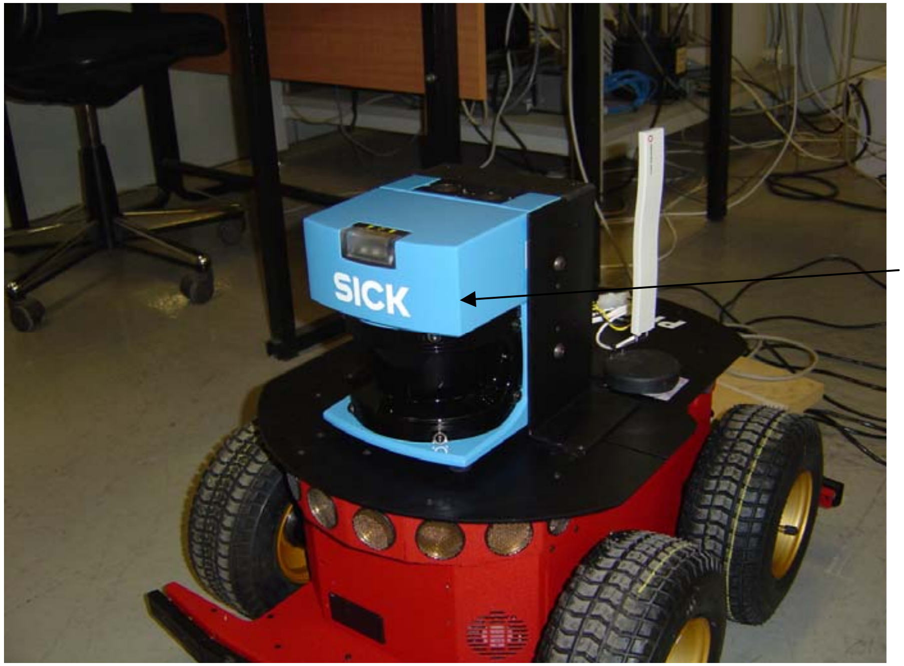
Further Principles

- So far only attention movement control executive function
- Must extend to motor response control (motor attention)
- Need prediction of next state after motor action (Forward model FM)
- Associated IMC & WM to achieve ‘thinking in one’s head’ – consider under reasoning
- Also emotion bias -> later

4. PLATFORMS/ALGORITHMS

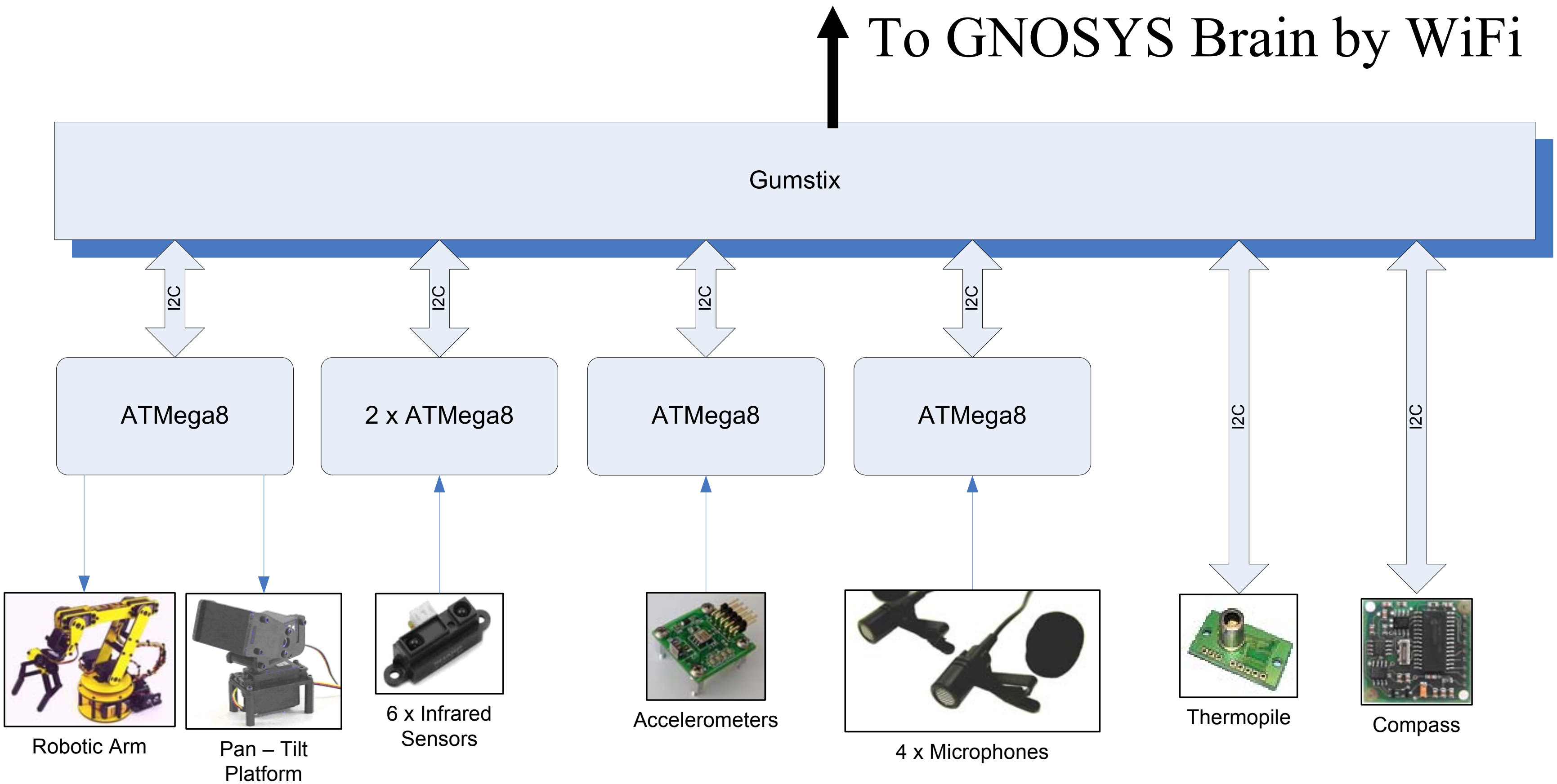
- Embodiment? – as robot or s’ware agent?
- But have covert attention (the only factuly that moves) => need have no embodiment & only sensory inputs (to attend to)
- For building response => embodied
- Can use wheeled robot platform
- Or up to humanoid (but many motor problems present outside cognition)

The SICK LMS 200 laser rangefinder attached to the basic robot platform (ZENON) (Pioneer P3AT robot)



Resolution of 0.5 Degrees, & scan rate of >10Hz

Sensing & Action on Gumstix Board (ZENON)

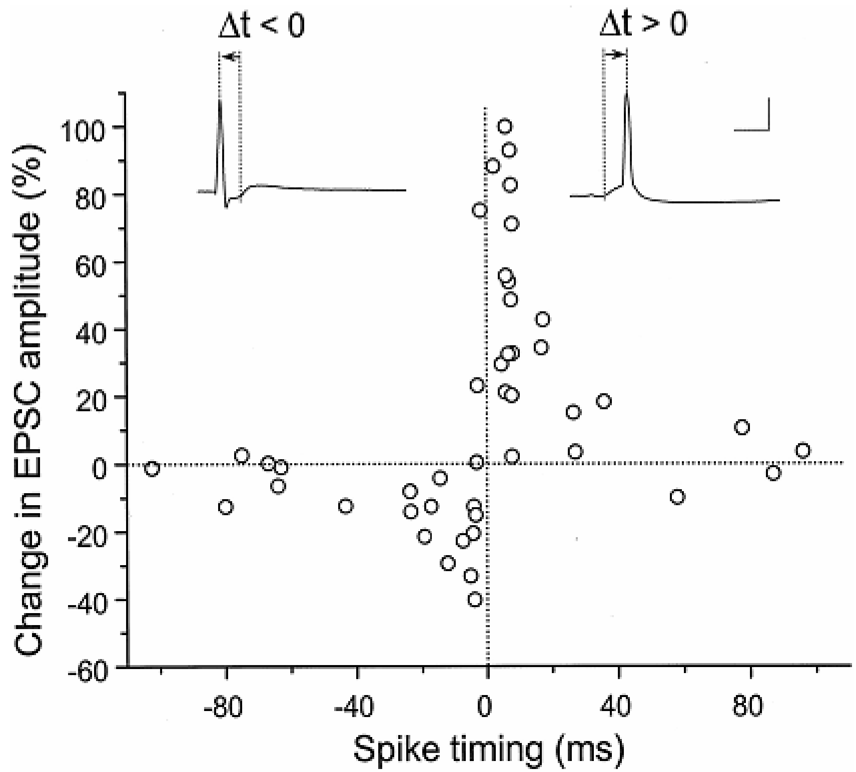


Algorithms

- Neurons as graded (eg simulink) or spiking (LIF/many channels/complex geometry, increasing computation load?)
- Graded ≈ mean field approx to spiking
- => graded better for coarse rep of space
- But difficult to use causal Hebbian (STDP) with graded
- STDP => sensitive to time so need spiking

STDP – Spike Time Dependent Plasticity (Bi & Poo, 98)

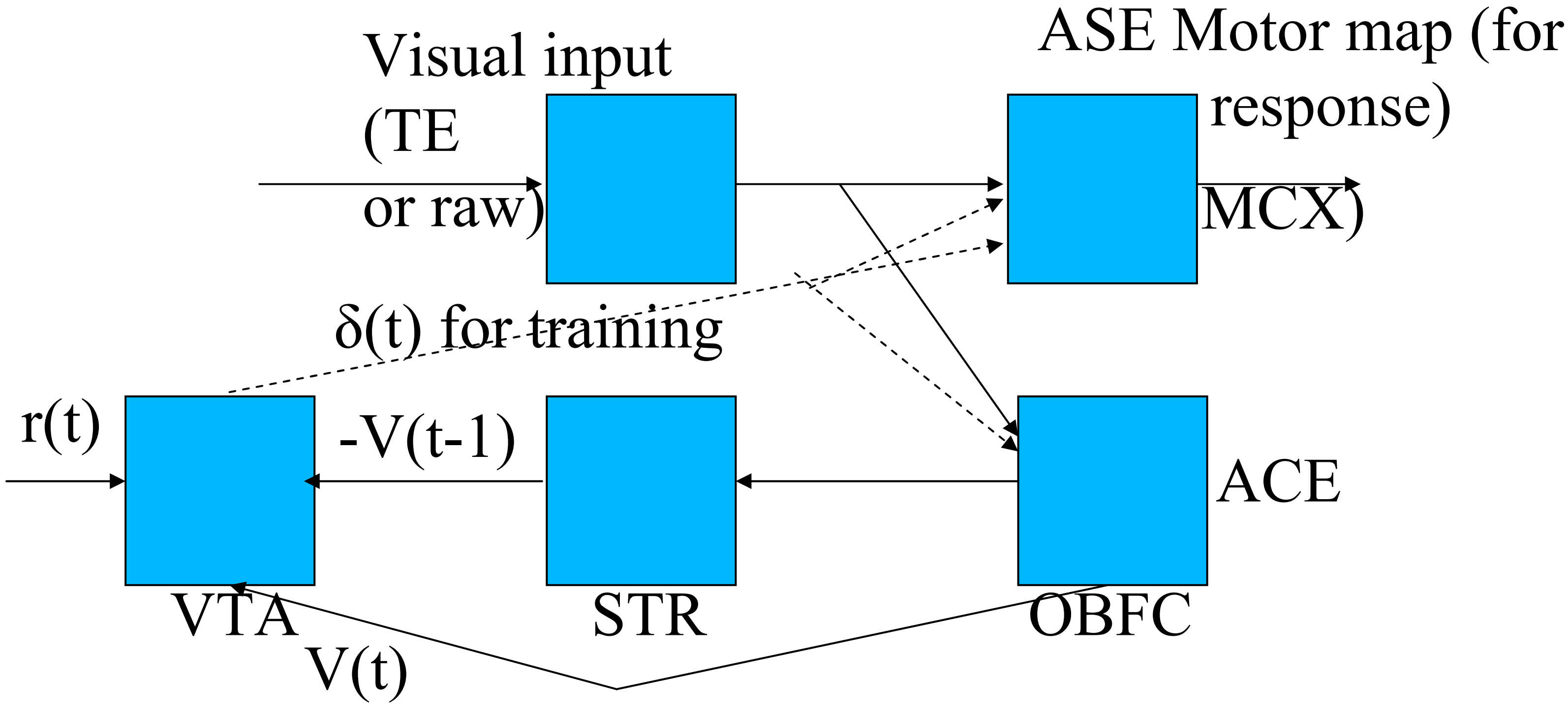
- Plasticity is critically dependent on interspike interval
- Positive interval => potentiation
- Negative interval => depression
- Model as 2-cpt process in synapse (MH/NT/JGT, 04)



Reward prediction P and reward prediction error δ

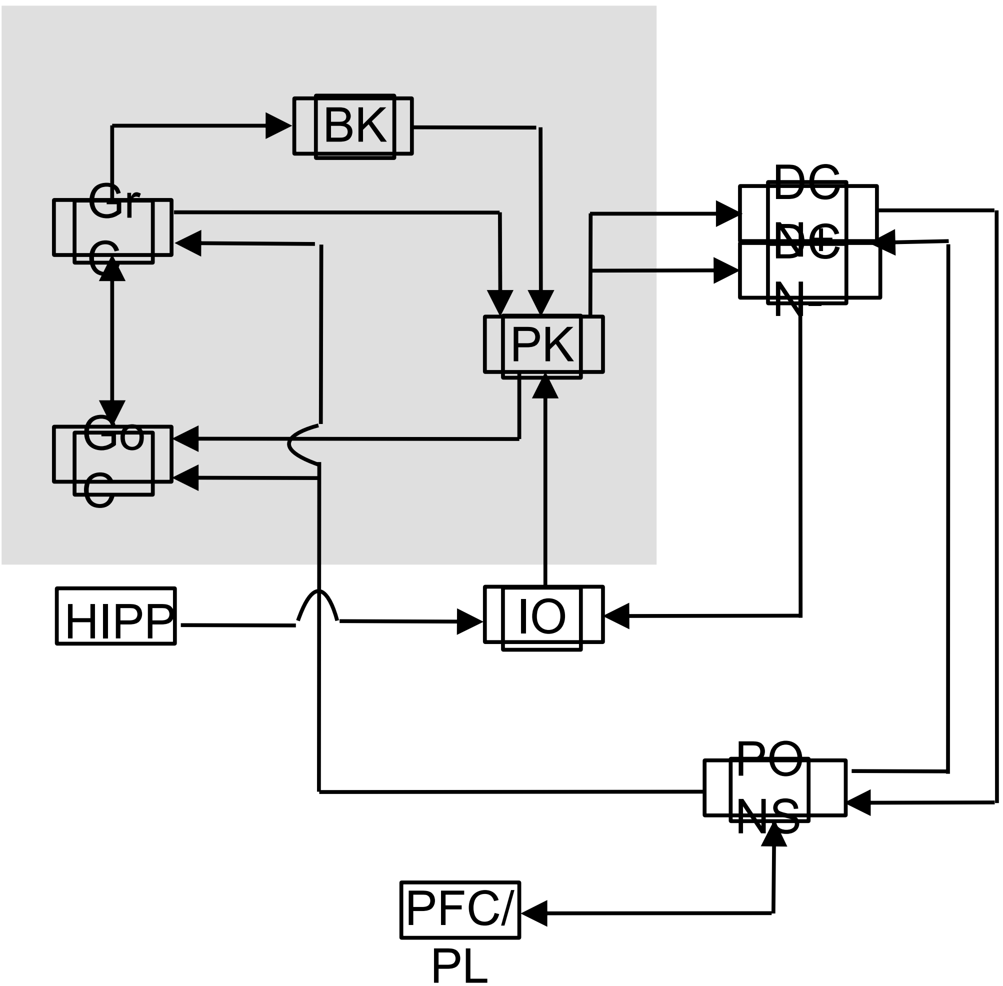
- Brain imaging experiments => δ coded in NAcc, and P in OBFC/Amygdala
- Amygdala coding of P rapidly updated
- OBFC coding of P not erased, but extended by context
- Use P for decisions among responses, and δ for learning to get P & responses optimally rewarded
- Coded in Netsim and runs fast

Value map architecture (KCL)



$\delta(t) = r(t) + V(t) – V(t\text{-}1)$: Learning by lrn rate = δ (as reward error predictor)

Error learning: Cerebellar Structure (& Associated Regions)



GrC granule cells
GoC golgi cells
BK basket cells
PK purkinje cells

DCN deep cerebellar nuclei (excit. & inhib.)
IO inferior olive
PONS pontine nuclei
HIPP hippocampus
PFC pre-frontal cortex

 inhibitory

 excitatory

Cerebellar Learning (BICS04: NT/MH/JGT)

- GrC->PK (parallel fibres) LTP/LTD (Ito)
  - LTP if GrC spike input without IO spike
  - LTD if GrC spike input follows IO spike <100ms
  - Feedback Error learning
  - $\Delta w_i = \partial_{LTD}\, GR_i\, CF(T)\, (\mathit{Ca\_conc} - 0.0175)$

    $+ \partial_{LTP}\, GR_i\, (1 - CF(T))\, (\mathit{Ca\_conc} - 0.0175)$

    where *Ca_conc* – calcium conc. in cell, GR_i is 1 if i^{th} GrC firing & 0 otherwise, CF(T) is 1 if IO spike arrived within last T ms & 0 otherwise (based on Medina et al)

Results

- Cb trained to insert ending in stem-ending generation process
- Assumed correct stem-ending pair stored elsewhere (hippocampal regions?) - teacher system
- Teacher used as US in conditioned learning paradigm, US1=stem, CS1=stem, CR1=stem
- Feedback incorrect CR1 to IO -> learns generate full stem-ending in Cb as outputs CR1 & CR2 from DCN
- This enables learning PFC rep. of stem-ending, such that contextual priming 'past' primes PFC/BG/THL rep. built by Cb now automatic

5. TASKS IN ANIMAL REASONING

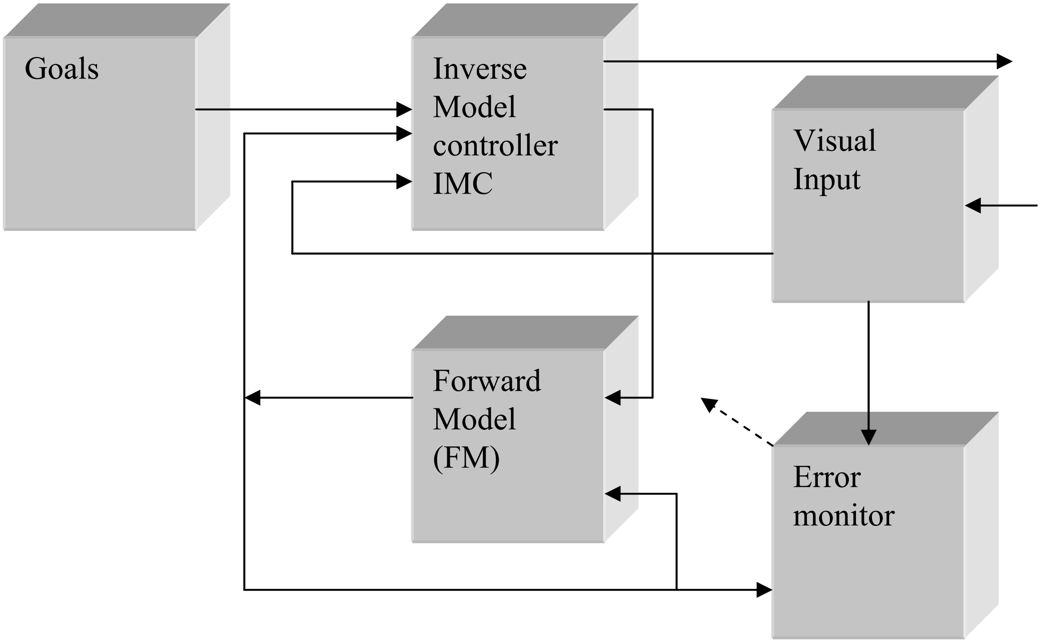
- Various non-linguistic reasoning tasks developed, extracted from tests on crows/chimpanzees/etc
- Associated architectures also proposed
- Need understand creativity (as in tool-making)
- Based on NN generalisation powers (natural semantic spreading)

Betty in action
(Betty = Caledonian Crow, Oxford)

(A Kacelnik & colleagues)

Success – now lunch!

Use of FM/IMC Pair in Reasoning



- This is a standard motor reasoning system, where the input to the IMC in the recurrent mode is solely that arising from the FM, with no input being used form the Visual Input module. The output (the dashed arrow) of the error monitor can be used to train the FM and IMC, using as input to the IMC the FM model input, but combined with the actual visual input when the latter is available. The output from the IMC is sent to the motor system (motor cortex, Cerebellum, basal ganglia, spinal chord) to move the muscles, and hence cause action on the stimulus generating the visual input. Need to relate to UGDIST to generate as automatic IMC/FM system.

Creativity (at L3)

- New tool-making
- Observed in Betty the crow
- Uses previously learnt models (of using a bent wire, and of bending a pipe cleaner)
- Had just previously used bent wire to achieve effect
- Now presented only with straight wire
- Semantic spreading/analogy ‘aha’ effect

Reasoning/Creativity Paradigms

- Numerous non-linguistic paradigms: chimp sticks to pull other sticks
- Betty use of bent stick
- Betty bending stick
- Stacking cylinders in each other (turn upside down) -> Tower of Hanoi
- Putting ball inside hollow cylinder then ask to find ball
- Pressing buttons
- Navigation (‘hide & seek’)

2-Sticks Reasoning Task (or button-touching to gain reward)

- Task: sticks S1 (short) S2 (long) gripper arm wants to reach food (or touch button) further than S1, closer than S2, but S2 only reached by S1
- I: Try gripper alone: NOGO => search
- II: See S1, try virtually => NOGO =>search
- III: See S2, ibid => OK => S2 rewarded as **subgoal** (new value map entry for S2)
- IV: S2 subgoal => try S1 to reach it => OK

Components needed

- FM/IMC pairs for virtual processing
- Adaptive reward/value map file of objects
- Buffer for actions and positions of objects
- Recognition of sticks and hard-wired grasp of sticks -> change parameters of gripper
- Ability to switch into virtual mode with no output of actions (but buffered internally)

Evaluation

- Evaluation criteria/test-cases to evaluate cognitive architecture
- Lower level: concept/ attention/ goal creation/ motor response/ learning attended motor response goals/ solutions as rules
- Higher level: reasoning paradigms
- Highest level: devise suitable patrolling tasks
- Criteria/metrics: Internal (separate module function); Behavioural (observed responses)

Psychological Paradigms/ Evaluation & Test Scenarios

- Type in commands GOTO ‘O’ (O = square, circle, triangle): observe response on trained GNOSYS, and internal activity (‘O’ activity on TE, etc)
- Type in commands GOTO ‘O’ (new O): observe learning processes and internal activity (as before)
- Extend to TOUCH, GRAB, RELEASE, MOVE ‘O’, STACK A on B
- Examine internal activity/accuracy during learning processes & trained responses, with reward basis + internal drives activated

Principles of Cognition

- Attention as highest control system (sensory and motor) using engineering control ideas (CODAM)
- Fuse motor + sensory by FM/IMC/WM triples
- Have lower level percept/concept reps
- Have STDP/Reward/Error learning->rules
- Have LTM by high capacity system (≈HC)

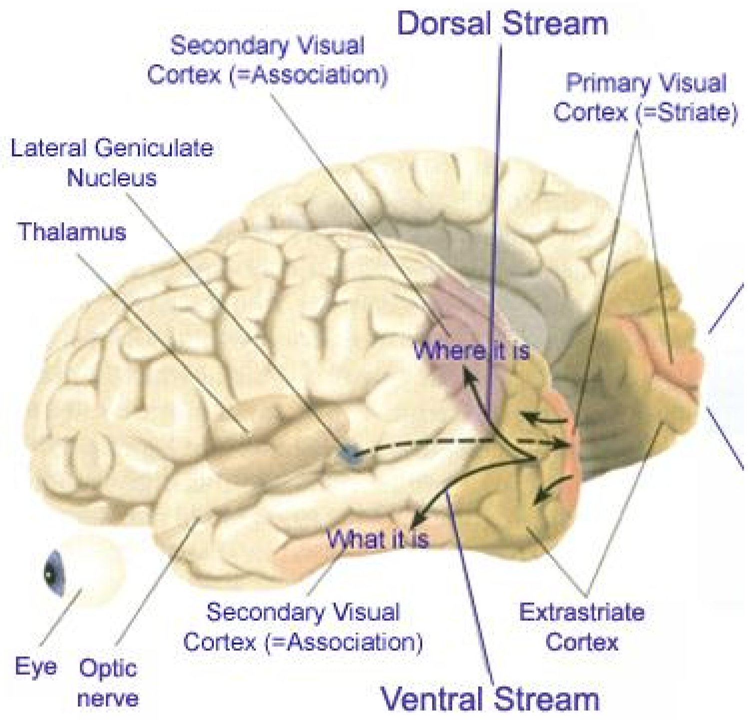
6. COGNITIVE ARCHITECTURES

- Based on executive control through attention transforming activities on WM buffers into desired activities
- Few models of these available
- Develop using CODAM with suitable extensions
- Need to extend to include emotional bias
- Need lower-level percept/concept creation

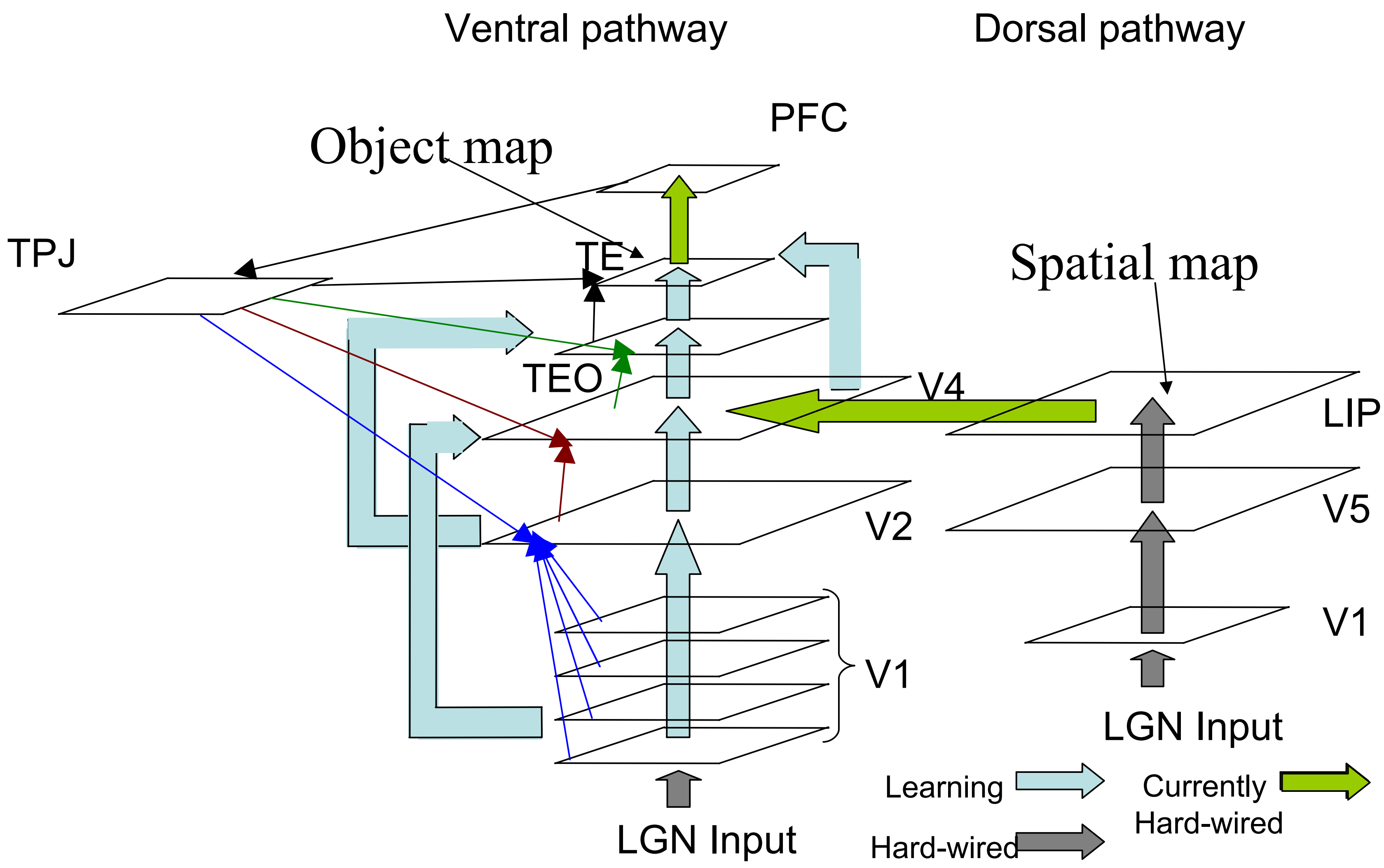
KCL: Percept/Concept/Attend System (MH/NT/CP/JGT)

- Hierarchical feed-forward learning of stimuli
- Created attention feedback (h’wired + l’rnt)
- Can learn new object representation L3)
- Present system trained on 56-node cluster
- To run on single processor, created ‘quick & dirty’ code
- Can run 28x28 LGN input in real time by qnd code on dual 3GHz PC
- Even faster code & more sensitive (8 orientat) + key-point picking methods (Huebner/Mallot)

Brain basis of Dorsal/Ventral Routes in Vision

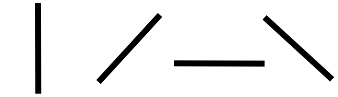


GNOSYS/EPSRC/BBSRC: Ventral & Dorsal Concept Learning

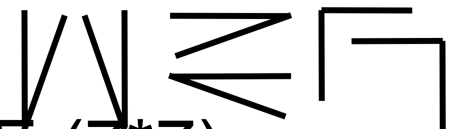


Architecture Details: Percepts KCL+C Pantev (IBS)

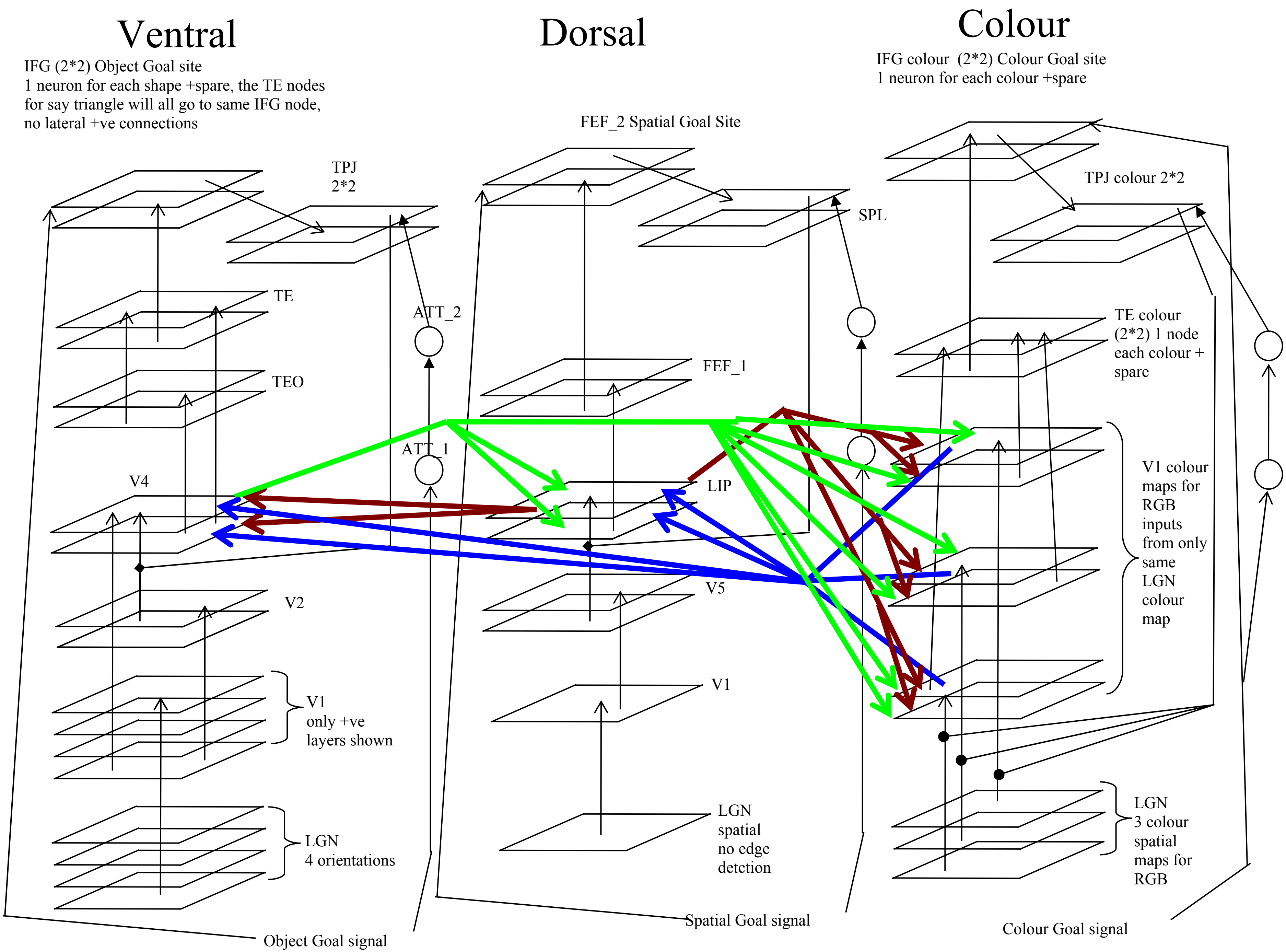
- V1: 4 excitatory & inhibitory layers for bar orientations, hardwired (28*28)



- V2 (50*50) trained on total set of pairs of bars (28), # start positions on retina = 121

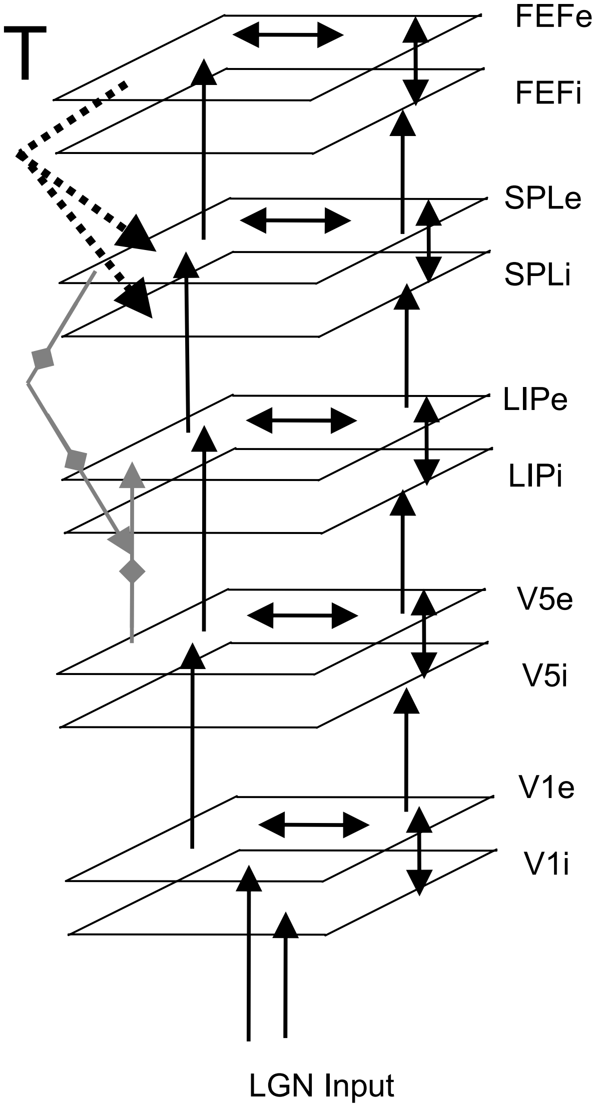


- V4 (28*28)->TEO (28*28/14*14)->TE (7*7) trained on arcs of circle, on triangle & square (121 start positions)
- Use 56 node Beowulf cluster to get weight files
- DL/VLPFC as goals-> attention feedback by TPJ
- Colour by 3 spatial maps in V1,TE


Ventral
Dorsal
Colour
IFG (2*2) Object Goal site
1 neuron for each shape +spare, the TE nodes
for say triangle will all go to same IFG node,
no lateral +ve connections
FEF_2 Spatial Goal Site
IFG colour (2*2) Colour Goal site
1 neuron for each colour +spare
TPJ
2*2
SPL
TPJ colour 2*2
TE
ATT_2
TEO
FEF_1
TE colour
(2*2) 1 node
each colour +
spare
ATT_1
V4
LIP
V1 colour
maps for
RGB
inputs
from only
same
LGN
colour
map
V5
V2
V1
V1
only +ve
layers shown
LGN
4 orientations
LGN
spatial
no edge
detction
LGN
3 colour
spatial
maps for
RGB
Object Goal signal
Spatial Goal signal
Colour Goal signal

Dorsal Attention System (NT/MH/JGT/KC/CP)

- Uses sigma-pi feedback) (NT/JGT
- Shows dynamics of guided attention feedback
- Have attended stimuli reps amplified in lower CX areas (KC/CP/JGT/NT)
- Develop now through cluster computing (KC/CP)
- Creates object -> position simulation



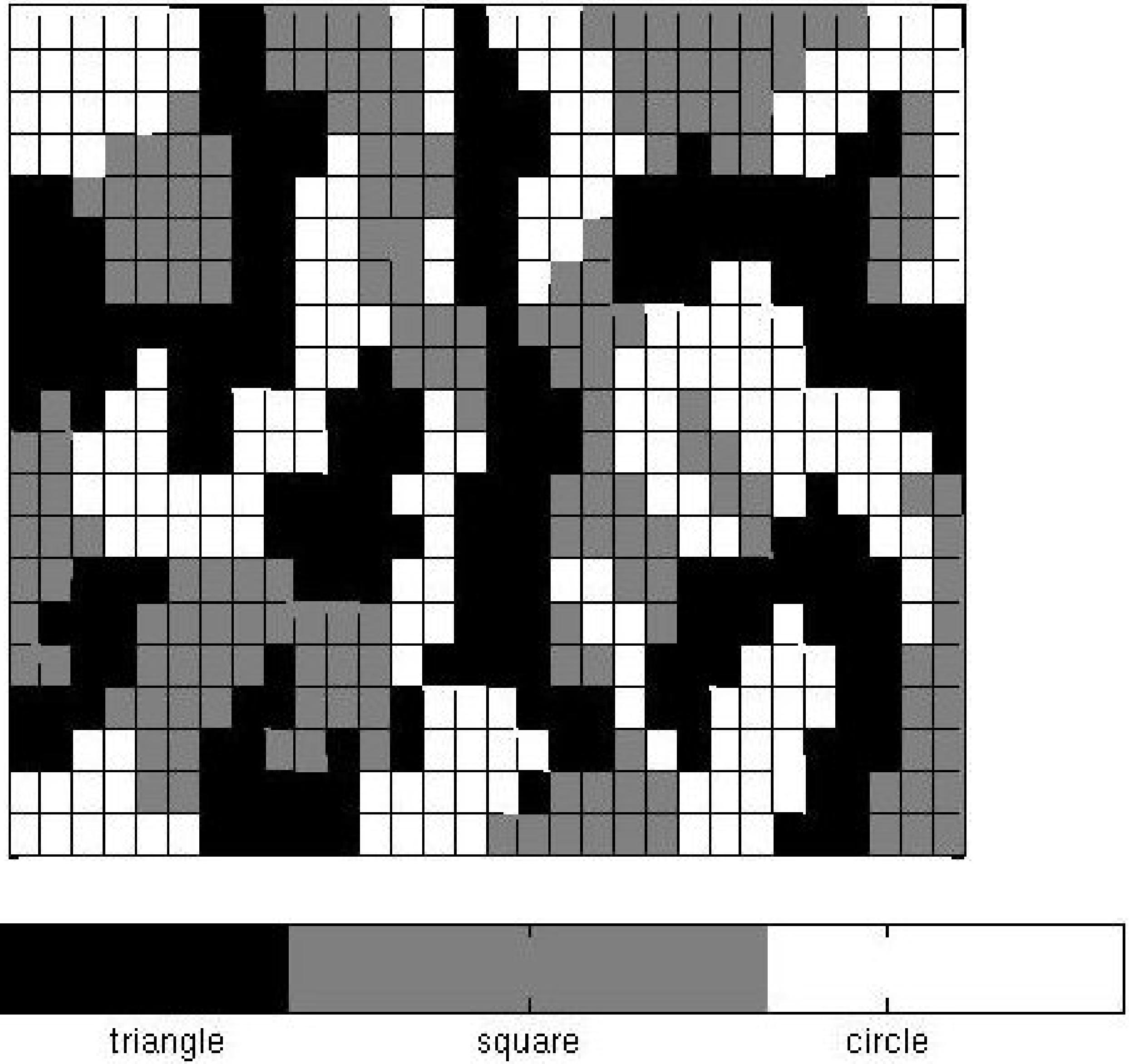
Activation of Sigma-pi sites

- All Sigma-pi sites have inputs via ATT_1 & ATT_2 nodes
  - ATT_2 nodes have a spontaneous firing rate that keep their Attention site (TPJ, SPL, TPJ_colour) inhibited unless appropriate external goal signal
  - This signal excites ATT_1 nodes which then inhibit the ATT_2 nodes, allowing for dis-inhibition of connected attentional site
  - Single or combinations of goals can bet set-up, ie. Red square, red postion, etc..

Lateral connections between streams

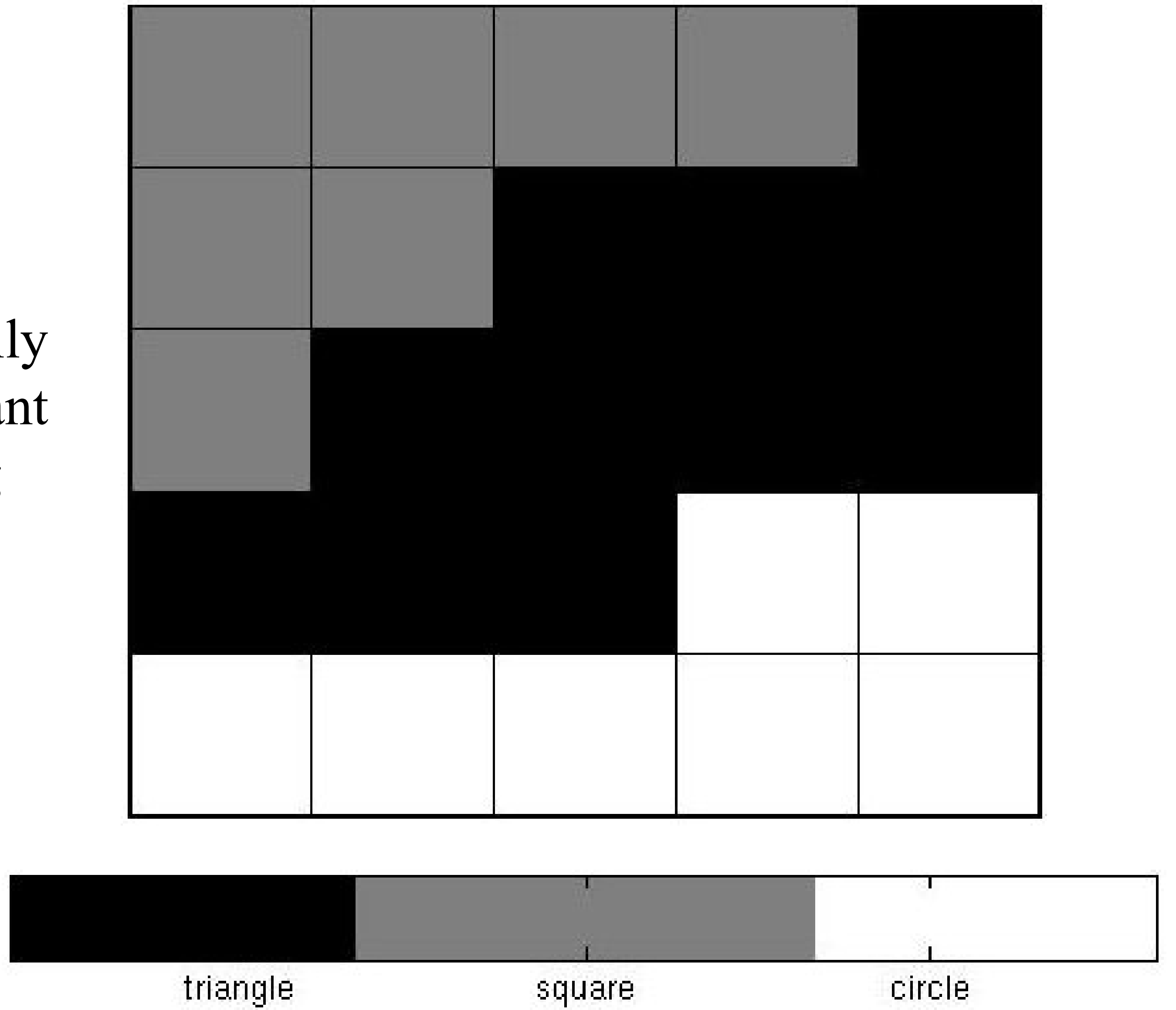
- Connections between the streams connect completely V3, LIP and V1_colour maps, from the excitatory layer to the +ve & -ve layers of the target.
  - This allows the attended stream to modify the other streams activation such that:
    - If attending colour the postion & shape can be obtained
    - If attending object the postion and colour can be extracted
    - If attending position the object and shape can be extracted
    - This is why there should be spatial information as this level of the streams

TEO Preference Map

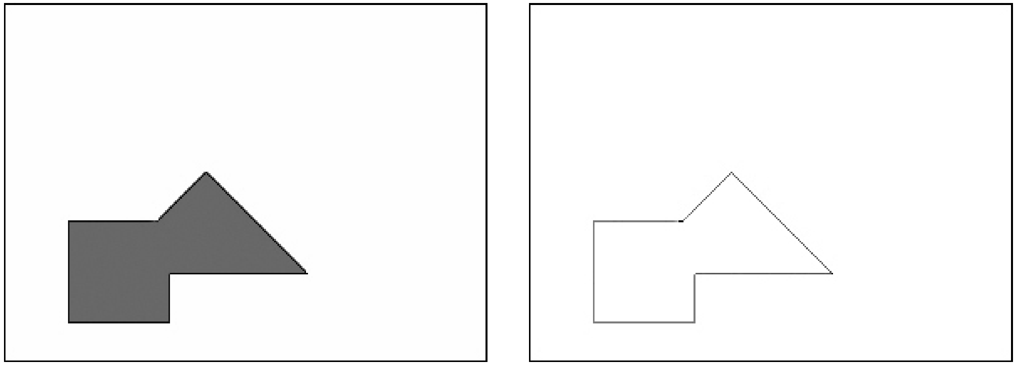


TE Preference Map

Spatially invariant coding

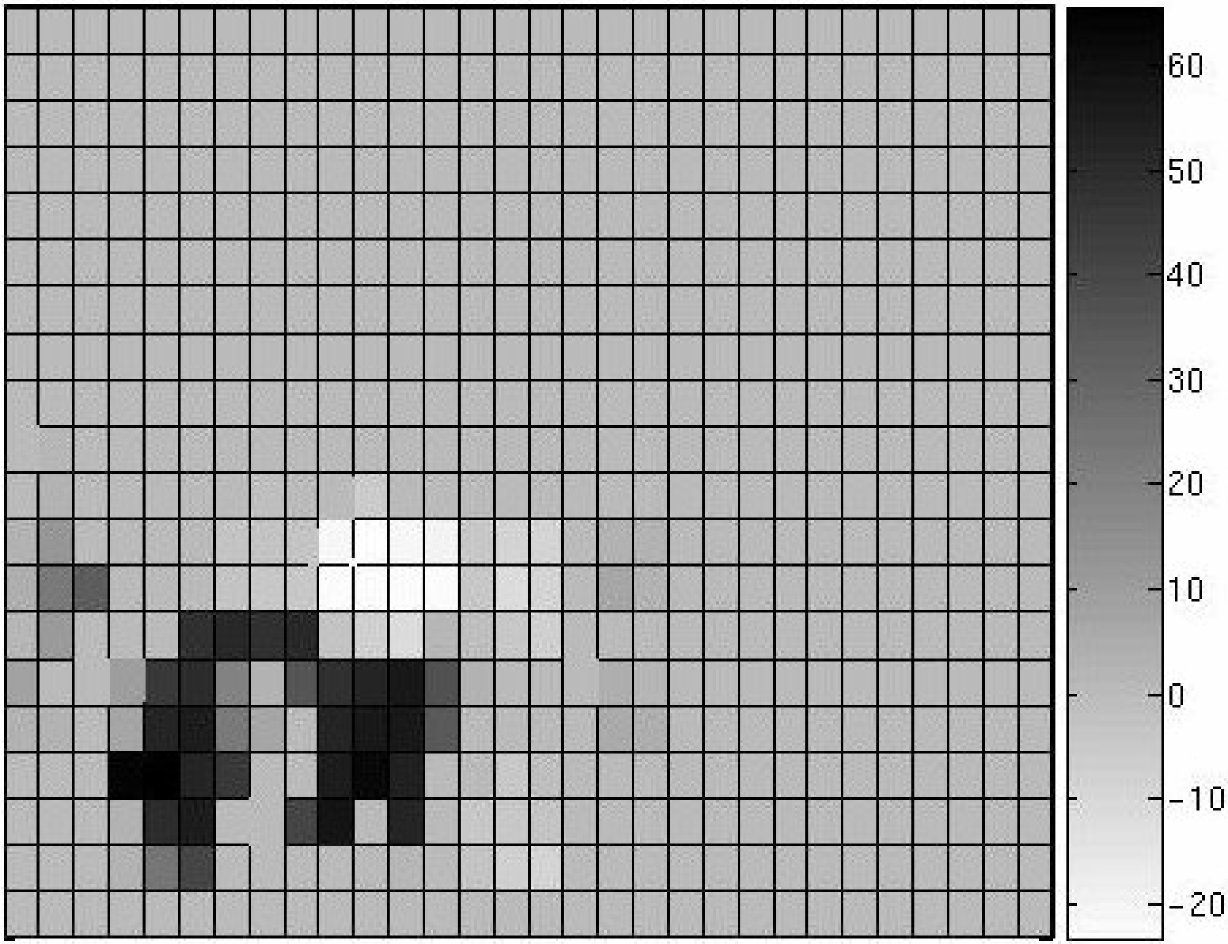


Occlusion: square & triangle (difficult occlusion in 2-d) (+NT/CP/MH)

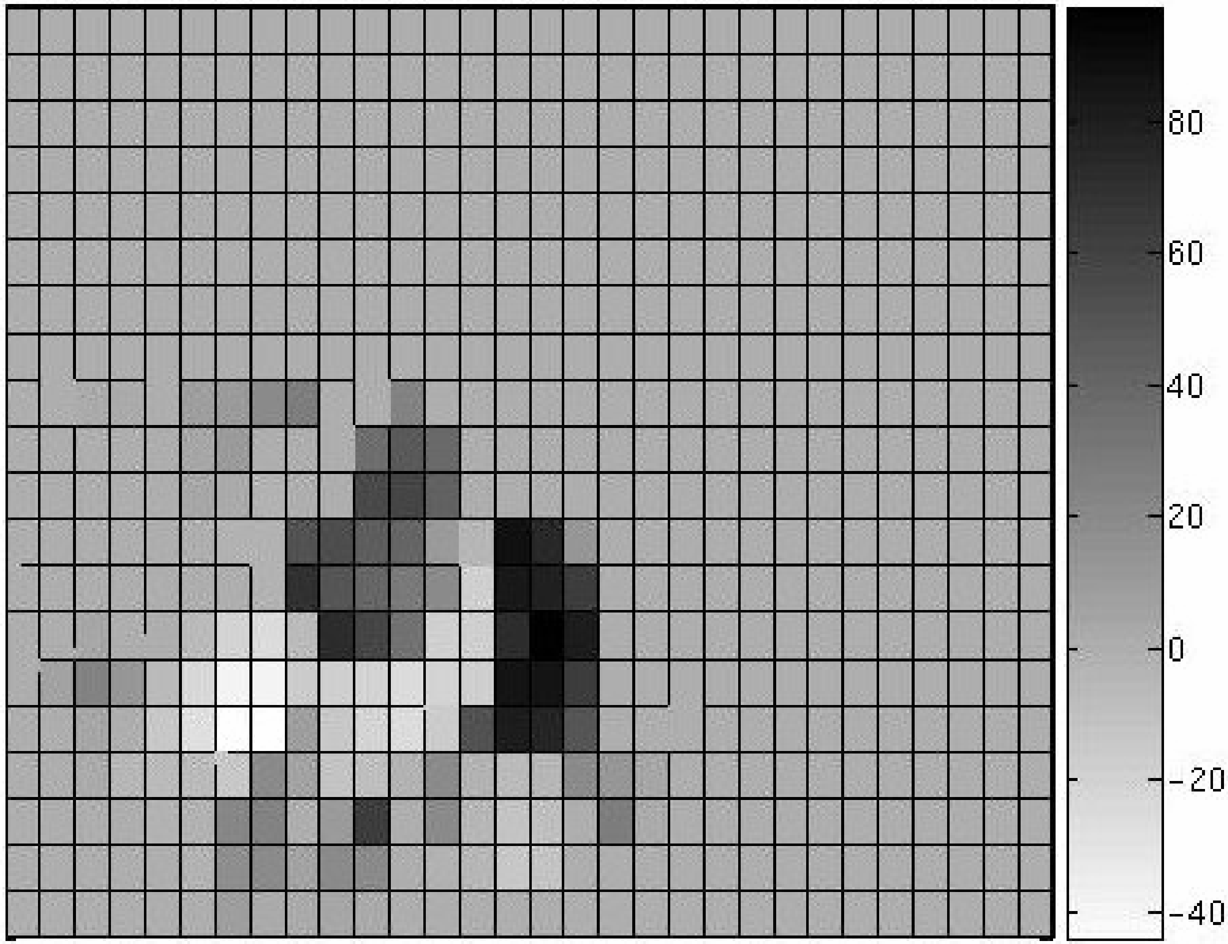


Full figures In silhouette

Attending to the Square in FEF



Attending to the Triangle in FEF



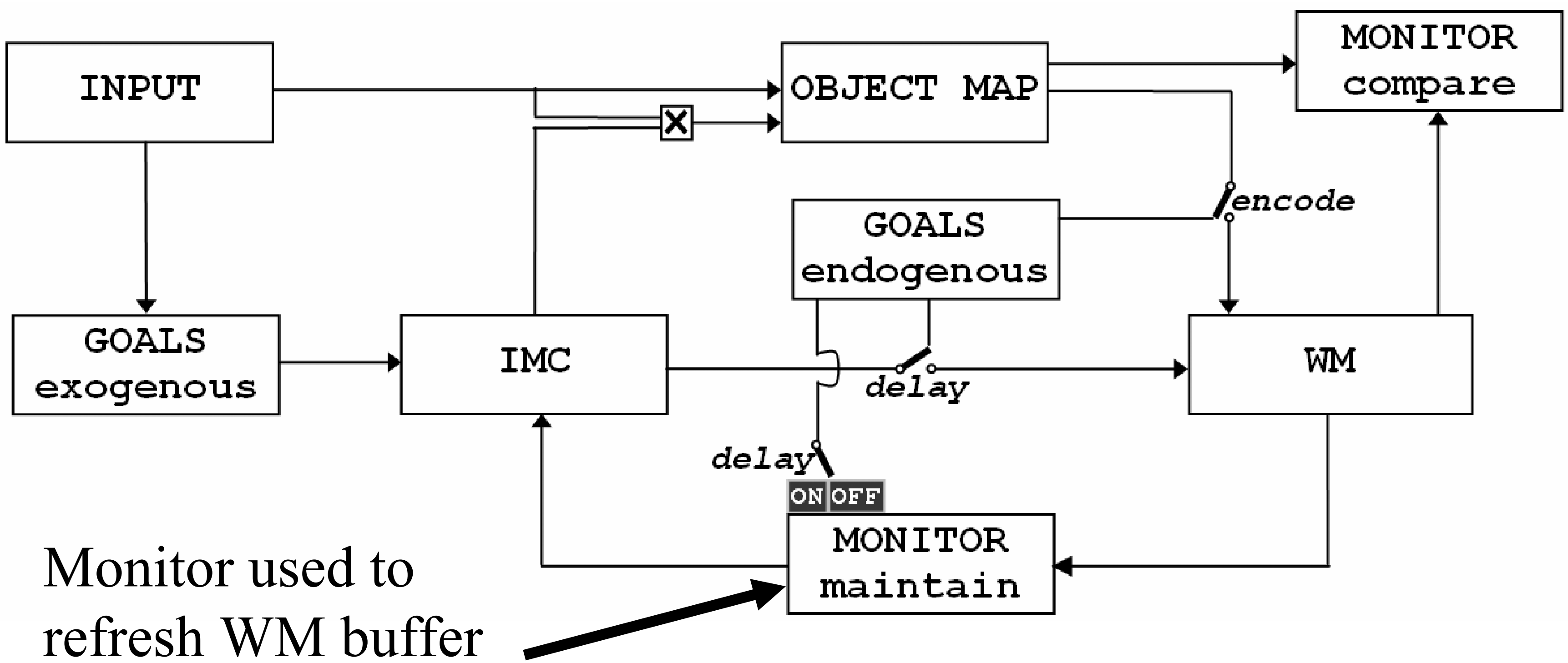
Summary

- Main effort on perception system + goals + attention (up to 3-d structures)
- Further update of perception system ongoing
- WM work ongoing
- Reward/value v attention work ongoing
- Architecture for attention/reward ongoing
- Architecture for executive functions being tested
- Beginnings of simple reasoning (<->MATHESIS)

Working Memory Control

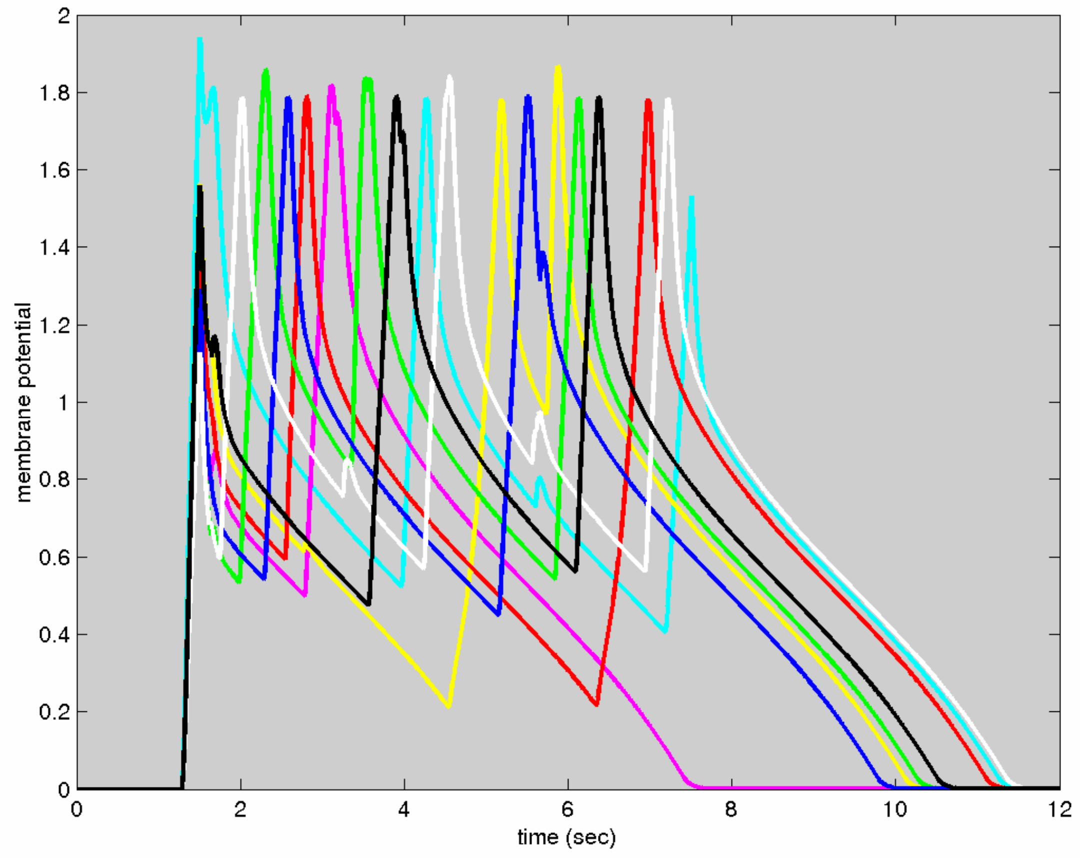
- Attention movement heavily intertwined with executive function (Nobre et al)
- Begin with CODAM basis to simulate:
- a) rehearsal; b) retrieval of probe
- Many models of buffer (continued) activity: bifurcation/two-state/attractor/…
- Few models of rehearsal process
- Try as attention redirected to representation (Awh et al, Nobre et al, Postle et al)

Attention/WM Control in Delay (ICANN06: JGT/NK/NF)

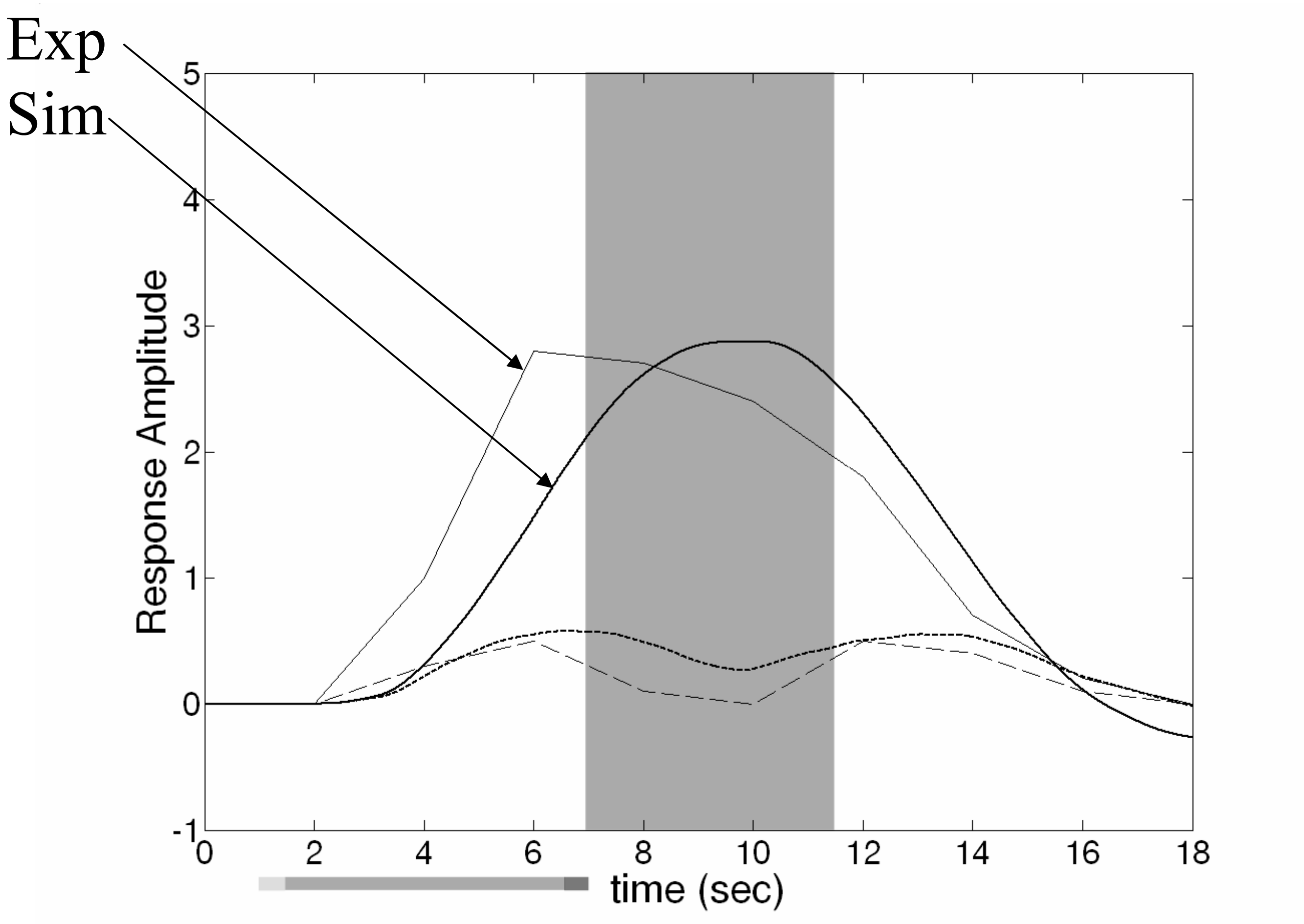


Monitor used to refresh WM buffer to keep activity above minimum level.

Maintenance/refresh Process of WM Nodes

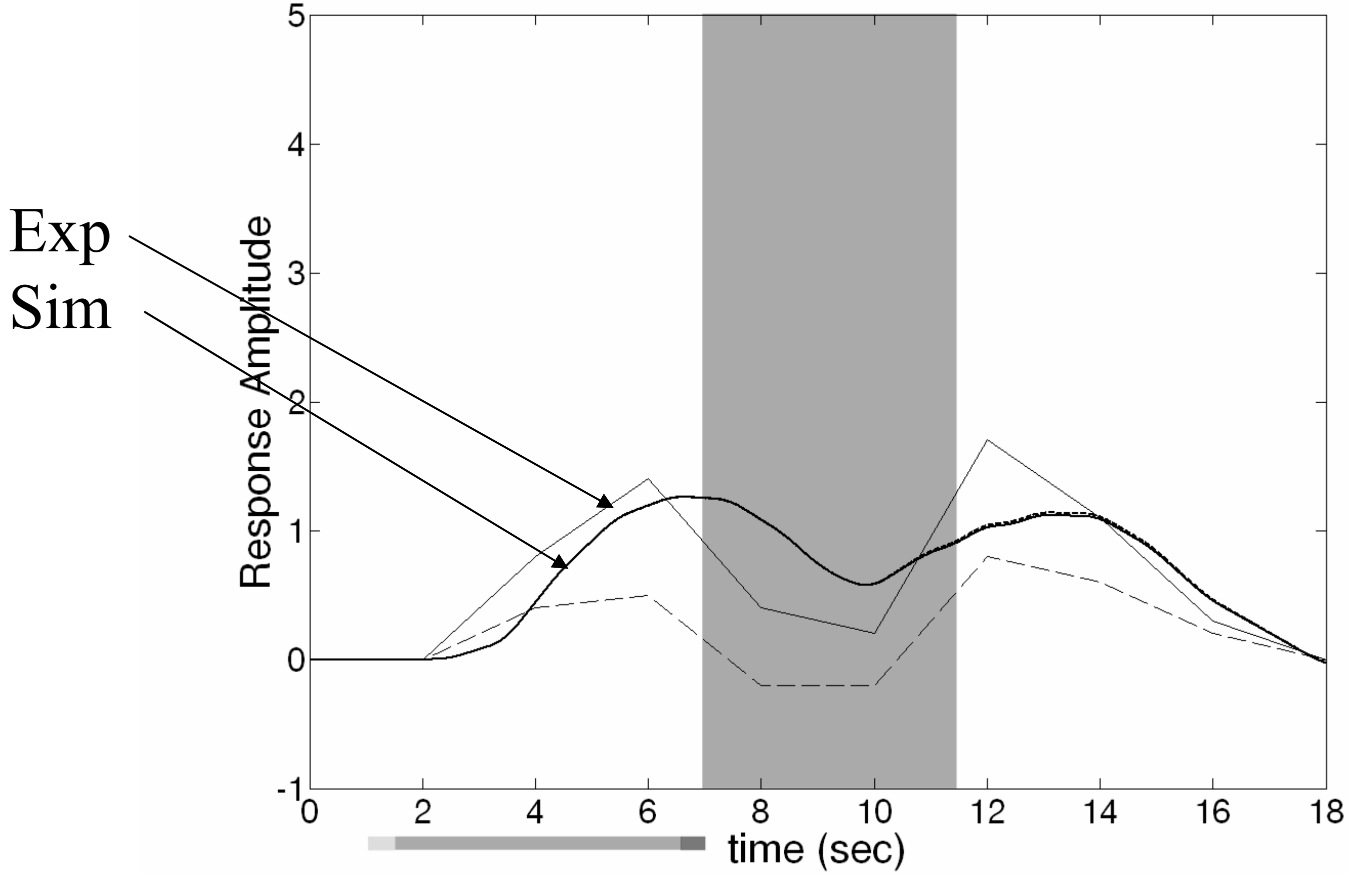


Simulated BOLD responses (bold lines) from model's IMC and the MONITOR/Maintain against the experimental BOLD signals (light lines) from right IPS.

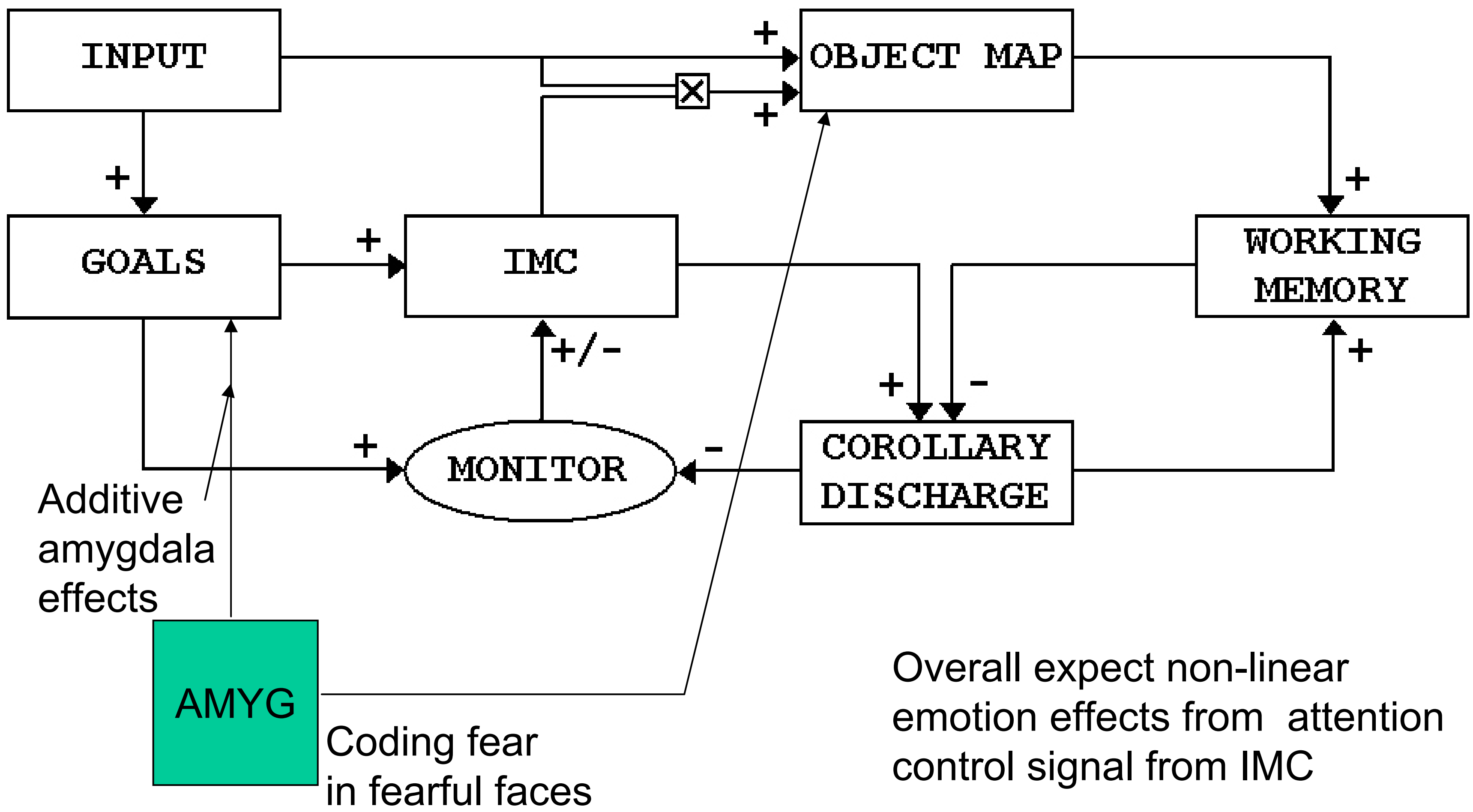


Pessoa et al
2002 data
Solid lines =
Correct trials
Dashed lines =
Incorrect trials

Simulated BOLD responses (bold lines) from model's Object Map plotted against experimental BOLD signals (light lines) from the DOC/IT



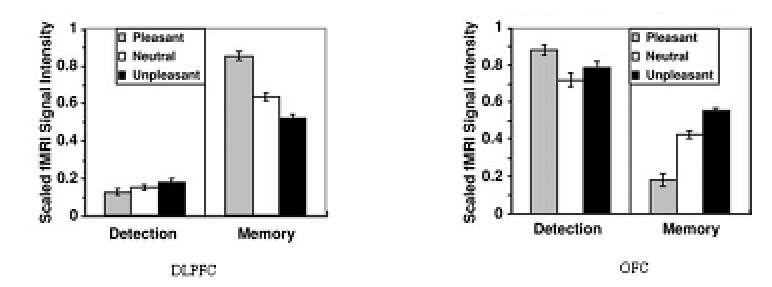
Extended CODAM model for value bias (->emotion)



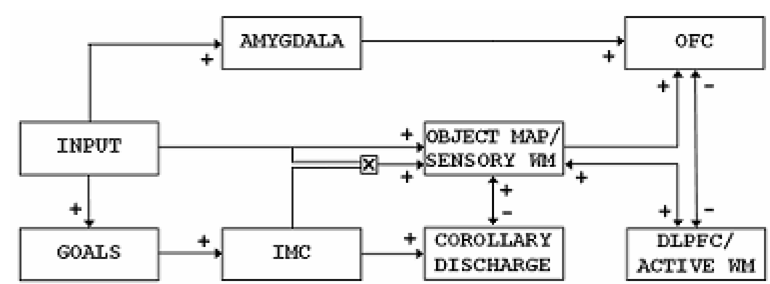
Overall expect non-linear emotion effects from attention control signal from IMC

KCL Attention/WM/Rewards (BICS2006: NK/NF/JGT)

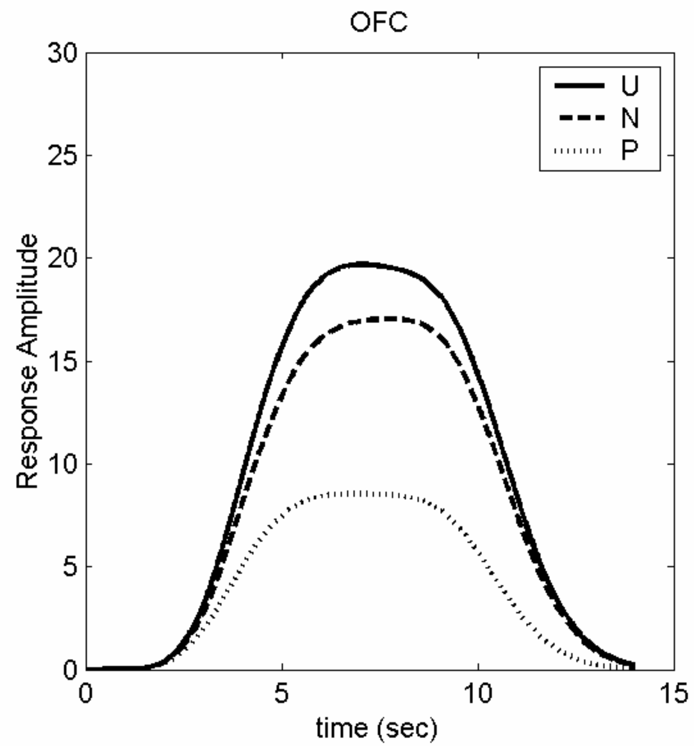
Perlstein imaging OFC<->DLPFC in WM delay and comparison:
In delay OFC<DLPFC
& P>N>U v P<N<U
Testing: OFC>DLPFC, all equal stimuli neutral/unpleasant/pleasant (= N/U/P)



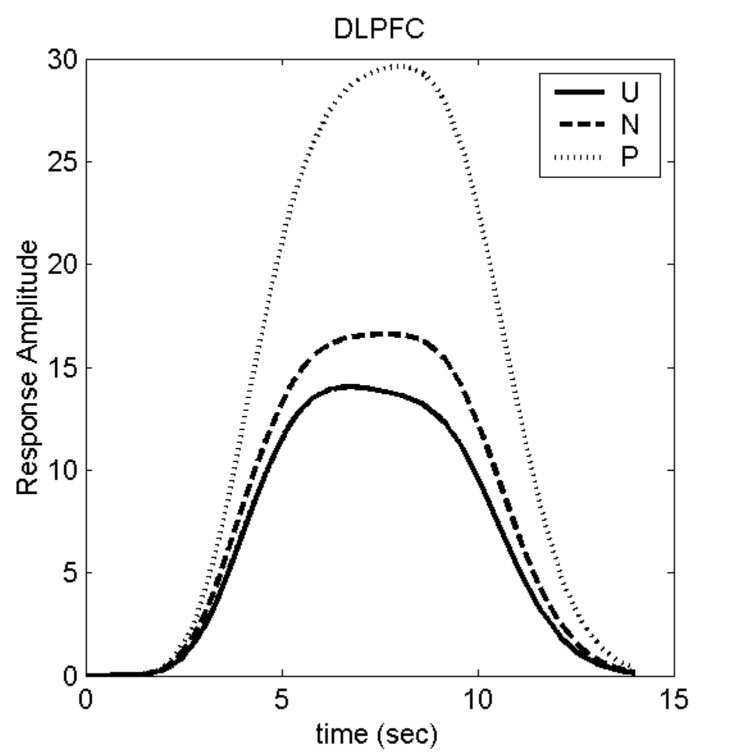
Use CODAM attention control model with extra bias from amygdala & OFC



Simulated BOLD Activation of OFC ≈ Experiment



Simulated BOLD Activation of DLPFC ≈ Experiment



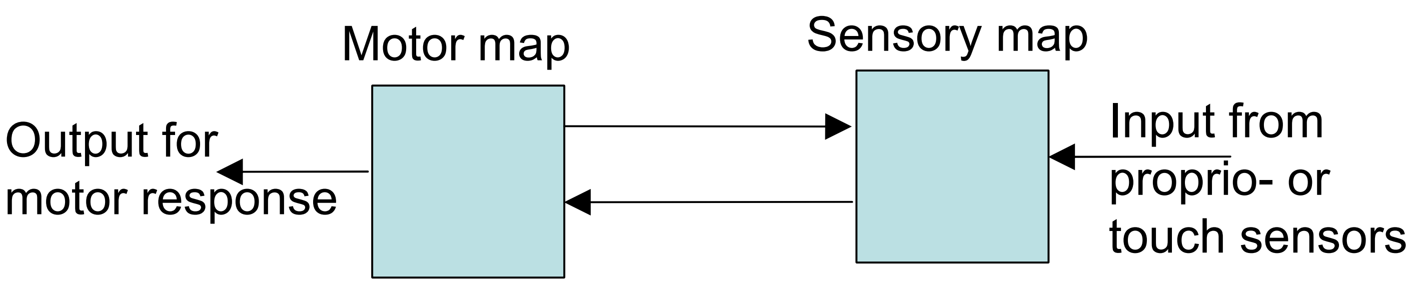
Plus predictions of more detailed temporal flow of activity in the various modules

Models of Motor Attention

- M1 map of motor actions to be taken by robot (move by motor/ manipulate by gripper)
- Need M1/S1 complex (for touch/ proprio-sensation): what feedback: to obtain?
- Construct M1 move map as simple reactive nodes by SOM for L2 (stacking, gripping, etc)
- Extend to sensory-motor control (attention)
- Gripper att'n map still to be constructed

Creation of Motor/Touch Maps

- Create by SOM (developed for V1->V2)
- Create topographic map of directions for movement (tuned cells)
- Create sensory topographic maps for robot (touch/proprio-sensation of gripper)

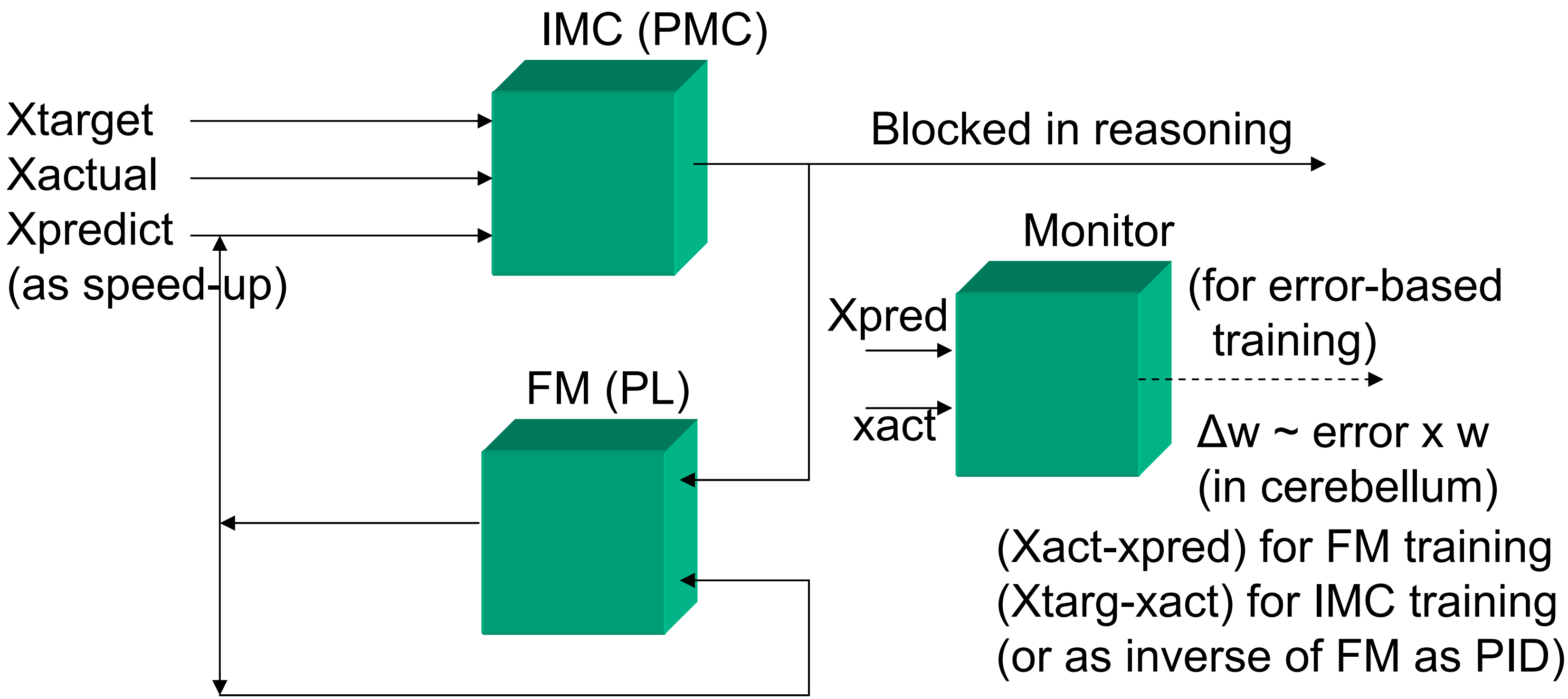


Internalisation of Tools

- Known to occur (car/tennis racket, etc become extension of body, arm, etc)
- Need to set up at software level:
  see stick -> grasp -> change length of gripper arm automatically
- Can then use for internal (virtual) reasoning as part of IMC/FM pairs (each has modification of parameters for lengths)

Creation of IMC/FM Pairs
(using recurrence to go through action-object sequence in a virtual manner)

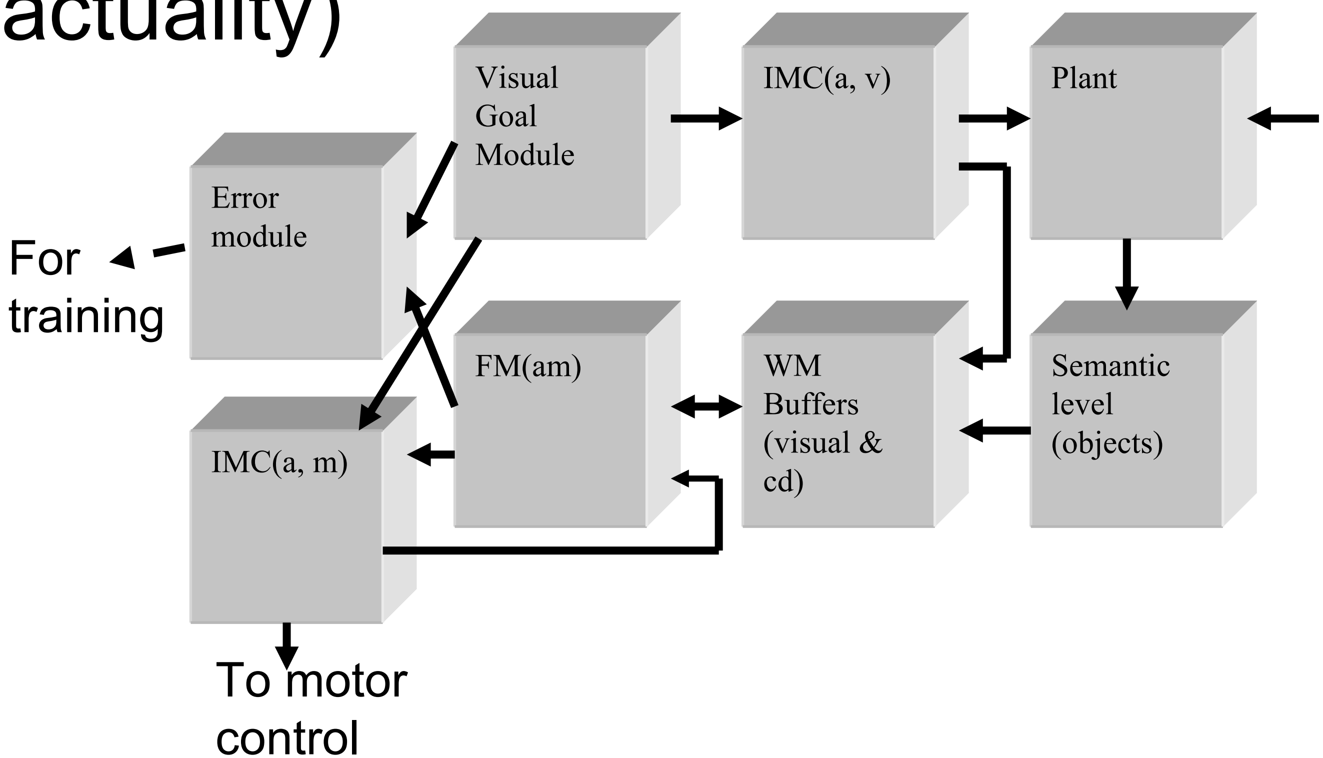
- Trains by internal error signal (in CG/Cb)



Cb model built and used
in TSSG as phrase inserter

Overall NN Architecture (KCL) (including attention control)

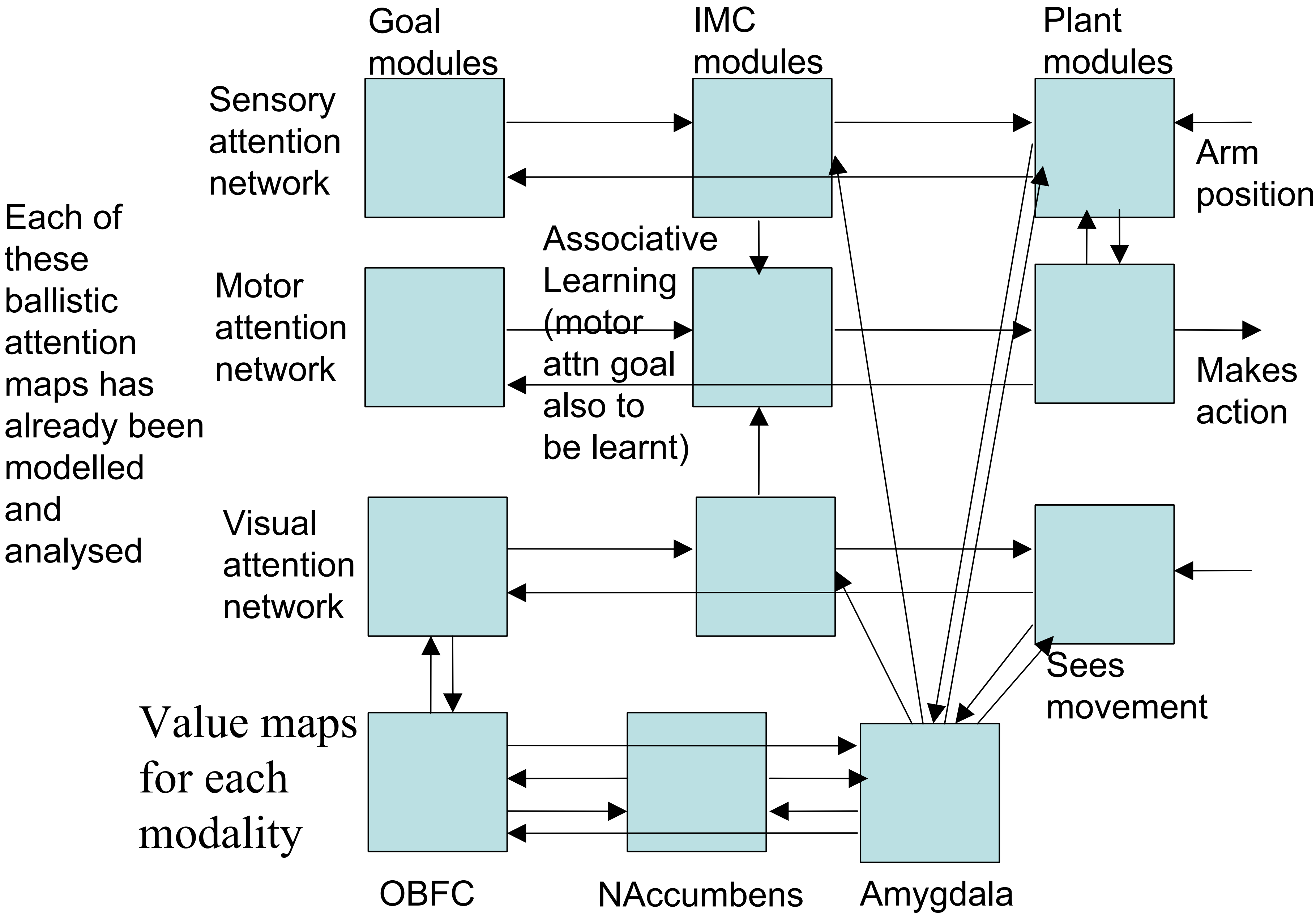
- Reasoning extended to being under attended recurrence (as occurs in actuality)



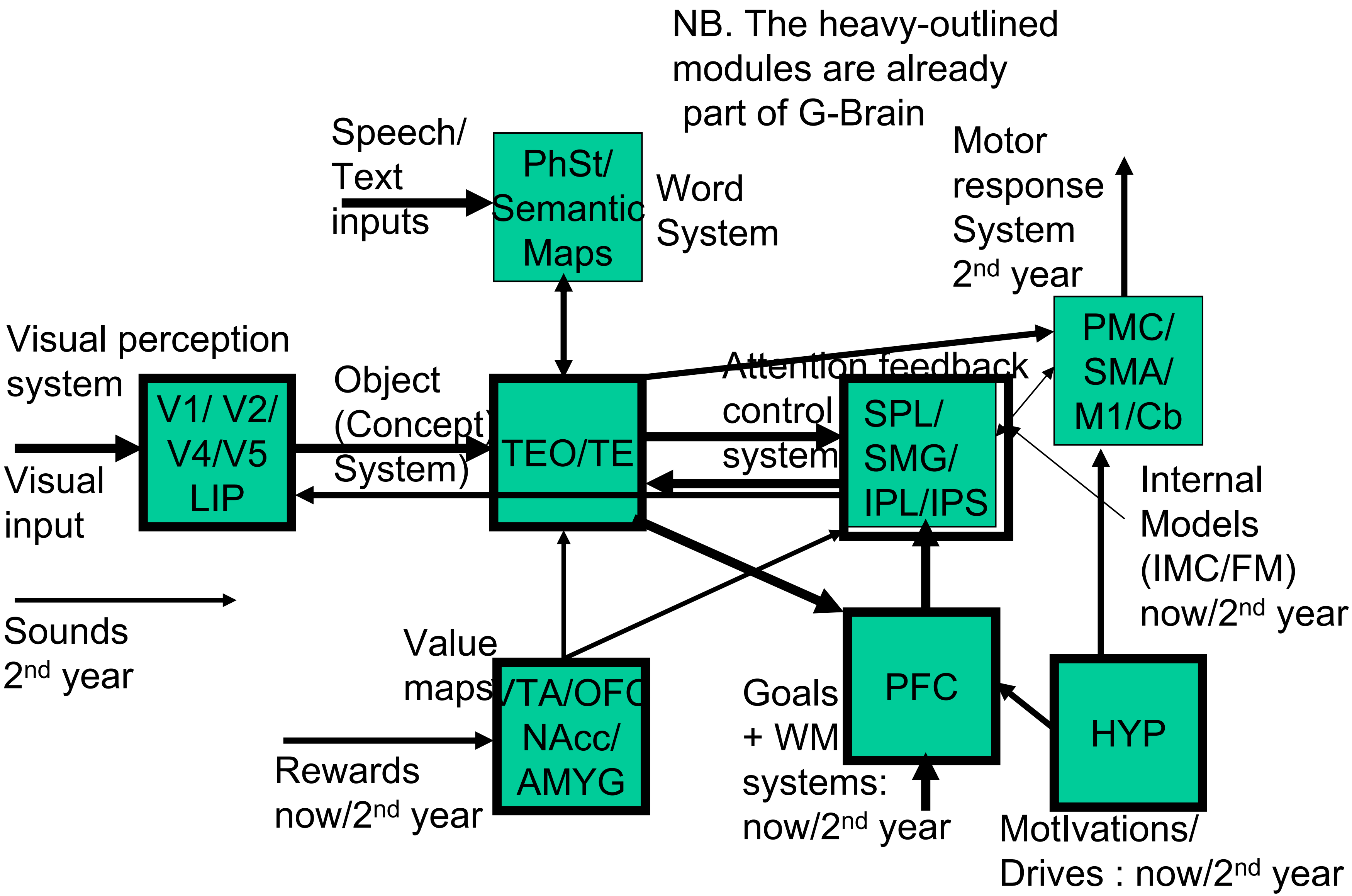
Concept System

- Composed of generalised perception representations = concepts
- Fuse with actions => action/object pairs (as actions possible on an object= more general ‘concepts’ = ‘OACs’ in JAST)
- Extending to TSSG through training recurrent systems under attention
- Much experience at KCL of ACTION net: as simplified model of frontal PFC/BG/TH
- Extening to include attention

HIGHEST LEVEL ARCHITECTURE



The Brain-Based G-Brain!



Where is Consciousness?

- CODAM model => consciousness by efference copy = corollary discharge of attention movement signal
- Gives ownership signal: ‘I am about to experience this content’ (by P300)
- Owner = ‘pre-reflective self’ (Husserl, Sartre, Western phenomenology) = PCE?
- Observable in N2 200-300 ms signal: just before P3 as fore-runner

7. Conclusions

- Results presented through analysis of activity of local modules
- Video of robot under control of G-Brain, at http://www.ics.forth.gr/gnosys/
- Results in CNS group at KCL CNS website http://www.kcl.ac.uk/research/cns/cns.html
- Progress in modelling animal reasoning
- Much more complexity needed at computational level
- Even then may need chip implementation to attain consciousness